

EVALUATING CORRECTIONAL PROGRAMS

*Edward J. Latessa, Ph.D.**

I. INTRODUCTION

Large sums of money are spent on correctional programs, and as a result legislators, funding sources, policy makers, and the public are demanding more accountability. Rigorous evaluation is no longer a luxury, it is a necessity. Programs and agencies are being asked, "Does it work?" and funding sources are not satisfied with anecdotal information and informal evaluations. Increasingly, evaluation research is becoming an essential part of a program's operation. This paper is designed to present an overview of some of the more common methods for evaluating correctional programs, and is designed to provide practitioners with a map for addressing the question, "Is a correctional program effective?"

Over the past twenty years I, along with my students and associates, have assessed over 500 programs throughout the United States, and we have conducted both large and small evaluations and studies that have helped us determine some of the significant factors that help comprise effective programs for offenders. We have learned that there are many attributes of effective programs that distinguish them from ineffective ones, and we have identified a number of areas that are important. These include: leadership and program development, staff, offender assessment, treatment, the use of core correctional practices, and quality assurance. This paper will examine some of these issues as well as the myths and constraints of evaluation research and ways in which correctional programs can be evaluated and assessed.

II. EVALUATION RESEARCH: MYTHS AND CONSTRAINTS

The purpose of evaluation research is to assist in rational decision-making, to provide feedback and monitoring of program performance, to add to our body of knowledge, and to help determine if a program is effective. It is important to emphasize that evaluation research should be considered a tool for improving program effectiveness. By making evaluation part of a program's operation, many of the myths and constraints confronting correctional research can be addressed. There are several correctional myths that should be dispelled:

- Some programs are not suitable for evaluation
- Programs can be informally evaluated
- Correctional knowledge advances by major breakthroughs
- That you can make comparisons across dissimilar programs
- Research costs too much
- There is a low payoff in research

When evaluating a program there are also a number of constraints that limit our ability to conduct correctional research. These include, but are not limited to: the political nature of the program, available resources, data, measurement techniques, and the identification of suitable comparison groups. Each of these issues should be taken into consideration when designing and conducting evaluation research in a program setting.

A. Political Nature of the Program

Many programs are political entities that emerge from the bargaining process. Performance is attached to

* Professor and Director of the School of Criminal Justice at the University of Cincinnati.

reputations, careers, jobs of staff, and expectations of policy makers and others. There are several steps that can be taken to reduce the effect of politics on an evaluation. First, make sure the program administrators fully understand the goals of the evaluation, and are kept informed as the study progresses. Second, help them view evaluation research as an on-going part of the programs' activities; preparing program staff for evaluation is an important step in the evaluation process. This will reduce the anxiety associated with program evaluation, and will make the staff feel more involved in the process.

B. Resources

A second constraint is the lack of financial and organizational support for evaluation research. Program evaluation is often conducted as an afterthought, and on a very limited budget. Making evaluation part of a program's operation and budget can go a long way to building a strong component for feedback and monitoring of program performance. Funding evaluation should be considered an investment for the long-term interests of the program.

C. Data

The condition, accessibility, and accuracy of correctional data are always a problem. Reliable and valid data are important if programs are to be accurately evaluated and assessed. One suggestion is to include the development of an automated database as part of an evaluation contract. This allows the evaluator to obtain data more quickly, (and at lower costs) while providing the program staff with a management information system that can track offenders over extended periods of time, and provide performance information without unnecessary delay.

D. Measurement

Measurement is always a problem in social science research. Tapping attitudes, knowledge, and behaviors is not an exact science. Recognize the limitations of data, and whenever possible use standardized instruments and scales.

E. Developing Comparison Groups

Finally, one of the biggest challenges in conducting evaluations is developing an adequate comparison group. While random assignment to treatment or control groups is considered the "gold" standard of research designs, it is often not possible when conducting real world research. As a result, researchers are often forced to use quasi-experimental design in which offenders not receiving the program are matched to the treatment group based on some relevant characteristics, the most important of which is a valid measure of risk.

F. Limitations of Evaluation Research

Evaluating the effectiveness of correctional programs is not easy even under the best of circumstances. First, political, ethical, and programmatic reasons may not permit random assignment of offenders to membership in the treatment or control group. Non-random assignment forces the evaluator to statistically make the groups comparable, an honored tradition in empirical research, but one that delivers results that are sometimes hard to communicate to policymakers and program directors (and sometimes to other researchers).

Even when random assignment designs are used there can be issues. Treatment or program effects can "bleed over" to the control group, or the intended treatment is inappropriately or unevenly applied. This makes it difficult to determine whether the treatment group members received needed treatment, and whether the control group remained "treatment free." After all, no program exists in a vacuum; historical accidents can affect either group, or one group more than another, or accidentally reinforce negative treatment effects in one group or another. For example, in our study of EPICS we randomly selected probation officers for training while the untrained became the control group. Even with this design, however, it is hard to imagine that there is no "contamination" since the trained officers work in the same office and often interact with their untrained colleagues.

Another major problem in evaluating treatment effectiveness in corrections is that it is rare to have only one treatment in operation at a time. For example, an offender ("Bob") may be sentenced initially to probation and restitution. The victim-offender interaction and mediation may have very positive effects on

Bob's attitudes and behavior. His drinking problem, however, may lead the probation officer to recommend that the court tighten the conditions of probation to include mandatory participation in substance abuse treatment, from which Bob derives some immediate and long-term benefit. Former antisocial friends may also become reacquainted with Bob, and misdemeanor crime may occur. Alerted by Bob's subsequent arrest, the PO may have Bob (a failure?) assigned to a group that includes relapse prevention techniques that assist him in identifying high-risk situations and coping with them. After three years of probation, when the victim's losses have been compensated, and with Bob securely employed in a job with a future and now voluntarily participating in substance abuse treatment, it is impossible to determine which of the treatment program elements will have been most effective in turning Bob around. Was it probation supervision, the quality of PO supervision, mediation and remorse associated with restitution, substance abuse treatment and direction, relapse prevention techniques, employment, or some combination of treatment elements? Because "probation" is a generic term that can refer to a combination of treatment, supervision, and intermediate sanctions, what element should be recognized as the "best intervention"? Finally, we need to deal with the question of whether Bob should be labeled a "success" or "failure" in corrections. Defining "failure" may mean using outcome indicators: arrest, reconviction or probation revocation, or incarceration (jail or prison). If the research design defines "success" as the absence of arrest, Bob failed: he was arrested. Yet the overall picture indicates that the arrest was just one critical incident in the long-range process of reintegration, one that Bob and his probation officer managed to overcome. Yet that single arrest incident in the three-year period would, from the perspective of reintegration, misclassify the probationer into the "failure" category. The bulk of evidence, however, clearly indicates that Bob was a success.

An example can be gleaned from a study we completed a number of years ago. One county in Ohio decided to create a new program for drug offenders on probation. The program was a day-reporting program and offenders were required to report every morning at 8 a.m. The program included acupuncture, drug testing, probation supervision, drug education, counseling, and support groups. The court wanted to know if the program was effective in reducing recidivism. Our initial design was to compare those offenders who went to this program to a risk-matched group of drug offenders who did not go to the program, but then the court wanted to know if acupuncture worked? To answer this question we had to create a study within the study. To isolate the effects of acupuncture, offenders in the program were randomly assigned to one of three groups: acupuncture — they received all of the program components; Placebo group — these offenders received acupuncture in the wrong spot and then got all of the programming; Control group — these offenders did not receive acupuncture, but did get the other program components. Only the acupuncturist and secretary (who kept the list) knew who was in the acupuncture and placebo groups. The staff did not know, nor did the offenders. Using a design like this allowed us to isolate the effects of acupuncture (which was not effective).

III. CORRECTIONAL EFFECTIVENESS

A. What is Effectiveness?

While the debate over correctional effectiveness will surely continue for some time, those attempting to evaluate and measure the worth of various strategies and programs found in corrections face a most difficult dilemma — defining "effectiveness." It is often said that even if recidivism is not reduced there are other benefits that can accrue from correctional programs. This may be true, especially for prison programs that can help address idleness, but like it or not, recidivism remains the primary measure by which we gauge the effectiveness of a correctional program. When asked if a correctional program "works," most do not care if the offenders or staff like the program, if participants feel better about themselves, or even if they completed the program. They want to know if the program helped change their behavior and if those who completed are less likely to recidivate than those that did not complete or go to the program.

There is little doubt that recidivism, no matter how it may be defined, should remain a main criterion; however, the need to measure additional outcome indicators appears obvious.

B. Recidivism as an Outcome Measure

Although most of the research we have examined focused on recidivism — as well it should — it is important to put this measure in perspective. Recidivism is and should be the primary outcome measure by which we assess correctional program effectiveness. However, recidivism is problematic for a number

of reasons. First, numerous definitions are applied, such as arrests, incarceration, technical violations, convictions, and so forth. How we define recidivism can determine the rate. For example, using a new arrest as the definition will result in higher recidivism rates than using return to prison. Second, the length of follow-up can be critical. For most offender groups a two- or three-year follow-up is sufficient; however, for some offenders, such as sex offenders, we need a much longer follow-up to adequately gauge recidivism. Third, recidivism rates can be influenced by both internal and external factors. For example, probation departments may change policies, such as increasing drug testing, which in turn can result in higher failure rates (internal), or police departments may focus on specific types of crimes such as random stops for drunk drivers (external). Finally, recidivism is often treated as a dichotomous variable: an all or nothing measure, when in fact we know that variations in this outcome measure exist. For example someone who is arrested for public intoxication is much less serious than someone arrested for armed robbery, but we often simply count them both as failures when examining program effectiveness. That said, recidivism is what is usually referred to when someone asks, "Does the program work?" This doesn't mean however, that we should not examine other "intermediate" measures. Let's now look at some ways this can be done.

C. Performance-Based Measures

In addition to long-term outcome measures, such as reductions in recidivism, we may also be interested in examining other intermediate measures. Unfortunately, in corrections we often count activities that have little or no relationship to the program or offender performance. An example would be counting the number of contacts between a probation officer and offenders. There is no empirical evidence that there is any relationship between the two factors, yet this is a common measuring stick in probation. The difference between counting an activity and a performance measure would be as follows:

- Activity: Counting the number of job referrals made
- Performance: Number of unemployed offenders at the time of arrest and the percentage employed within six months after being placed on probation

Often times documenting the performance of offenders will help you determine treatment effects. Examples would be:

- Reductions in dynamic risk/need assessment scores
- Changes in pre-/post-measures, such as improvement in test scores, changes in attitudes, behaviors, etc.
- Changes in problem areas (e.g. drug test results)
- Completion of behavioral objectives (meeting treatment plan)
- Substance Abuse: drug tests, attitude change, days of abstinence, etc.
- Education: improvement on standardized achievement tests
- Employment: days employed, earnings, savings, contributions to support, etc.
- Mental Health: days hospitalized (pre-/post-treatment)

By focusing on performance rather than activities, a correctional program can develop intermediate goals, which, if achieved, can serve as a prelude to reductions in recidivism. Osborne and Gaebler (1993) have identified seven principles for results-oriented management:

- I. What gets measured gets done;
- II. If you don't measure results, you can't tell success from failure;
- III. If you can't see success, you can't reward it;
- IV. If you can't reward success, you're probably rewarding failure;
- V. If you can't see success, you can't learn from it;
- VI. If you can't recognize failure, you can't correct it; and
- VII. If you can demonstrate results, you can win public support.

D. Measuring Program Quality

Few would argue that the quality of a correctional intervention program has no effect on outcome. Nonetheless, correctional researchers have largely ignored the measurement of program quality. Traditionally,

quality has been measured through process evaluations. This approach can provide useful information about a program's operations; however, these types of evaluations often lack the "quantifiability" of outcome studies. Previously, researchers' primary issue had been the development of criteria or indicators by which a correctional program can be measured. While traditional audits and accreditation processes are one step in this direction, thus far they have proven to be inadequate. For example, audits can be an important means to ensure that a program is meeting contractual obligations or a set of prescribed standards; however, these conditions may not have any relationship to effective intervention. It is also important to note that outcome studies and assessments of program quality are not necessarily mutually exclusive. Combining outcome indicators with assessments of program quality can provide a more complete picture of an intervention's effectiveness. Fortunately, there has been considerable progress in identifying the hallmarks of effective programs.

III. INTRODUCTION TO PROGRAM ASSESSMENT

A. Getting Inside the "Black Box" of a Correctional Program

The characteristics and quality of a correctional program often help determine its effectiveness. The question is "how do we get inside the 'black box' of a correctional program?" Let's now turn our attention to how we can measure the integrity of a correctional program.

Examining the "input" of a program is usually referred to as a process evaluation. Process evaluations usually involve a more qualitative methodology than an outcome evaluation. A process study helps determine whether the program is operating as designed. The problem of course is that a program may in fact be operating efficiently, but not effectively. For example, a drug education program may be doing a great job of teaching offenders about the harm that drugs can do to them, without being effective in reducing drug usage.

The other problem with traditional process studies is that they do not provide a "quantitative" measure. One way to think of this would be the example from offender assessment. Some assessment processes gather a great deal of information about the offender (i.e. criminal history, employment, drug use, family, education, etc.). The problem of course is that when they are done, they don't really have a good way to pull it all together to quantifiably measure risk. Now compare that to using an actuarial assessment tool. The same information is gathered, but when you are finished you produce a score that in turn helps tell you the probability of recidivism, as well as whether or not the offender scores "high", "medium" or "low" in each domain.

B. Correctional Program Assessment Tools

So how do we quantifiably measure program integrity, and what factors are examined? One tool used was developed by Gendreau and Andrews: the Correctional Program Assessment Inventory (CPAI). This instrument is based in part on the results from meta-analyses of correctional effectiveness studies. It is a tool for assessing correctional programs based on empirical criteria. However, unlike traditional process evaluations or audits that simply measure if you are doing what you say you are, this process looks at the degree to which a correctional program is meeting the principles of effective intervention. Over the years researchers from the University of Cincinnati have validated and tested items from the CPAI and created some new tools including the Correctional Program Checklist (CPC), the CPC for Prison Groups, CPC for Drug and Mental Health Courts, and the CPC for Community Supervision Agencies. These tools have been used to evaluate programs all over the United States and several other countries.

C. The Evidence-Based Correctional Program Checklist (CPC)

The CPC is a tool we developed for assessing correctional intervention programs and is used to ascertain how closely correctional programs meet known principles of effective intervention. There are several ways the CPC has been used:

- To evaluate the extent to which correctional treatment programs adhere to the principles of effective intervention
- To assist agencies with improving and developing the services provided to offender populations
- To evaluate funding proposals as well as external service contracts

- To assist in program development
- To stimulate research on the effectiveness of correctional treatment programs

Several recent studies conducted by the University of Cincinnati on both adult and juvenile programs were used to develop and validate the indicators on the CPC. These studies found strong correlations with outcome between both domain areas and individual items. We have also conducted over 500 program assessments across the USA and elsewhere and have developed a large database on correctional intervention programs.

Table 1 shows some of the types of programs that we have assessed.

Table 1. Types of Programs Assessed with the CPAI or CPC

Boot Camps
Cognitive Behavioral Programs
Community Correctional Facilities
Community Supervision Agencies
Correctional Education Programs
Day Reporting Centers
Diversion Programs
Drug Courts
Group Homes
Halfway Houses
Intensive Supervision Units
Institutional Sex Offender Programs
Institutional Treatment Programs
Jail Based Substance Abuse Programs
Outpatient Substance Abuse Programs
Prison Programs
Residential Correctional Programs for Parolees
Residential Correctional Programs for Women
Residential Substance Abuse Programs
Residential Substance Abuse Programs for Habitual Drunk Drivers
School Based Programs
Sex Offender Programs
Therapeutic Communities, both institutional and community based
Work Release Facilities

The CPC is divided into two basic areas: content and capacity. The capacity area is designed to measure whether a correctional program has the capability to deliver evidence-based interventions and services for offenders. There are three domains in the capacity area including: Leadership and Development, Staff, and Quality Assurance. The content area focuses on the substantive domains of Offender Assessment and Treatment, and the extent to which the program meets the principles of risk, need and responsivity. There are a total of seventy-seven indicators, worth up to 83 total points that are scored during the assessment. Each area and all domains are scored and rated as either "highly effective" (65% to 100%); "effective" (55% to 64%); "needs improvement" (46% to 54%); or "ineffective" (45% or less). The scores in all five domains are totaled and the same scale is used for the overall assessment score. It should be noted that not all of the five domains are given equal weight, and some items may be considered "not applicable," in which case they are not included in the scoring. Appendix A shows the items that are included in the CPC.

D. Information Gathering Process

Data are collected through structured interviews with selected program staff and program participants and observation of groups and services. In some instances surveys may also be used to gather additional

information. Other sources of information include policy and procedure manuals, schedules, treatment materials, manuals, and curricula, a review of a sample of case files and other selected program materials. Once the information is gathered and reviewed, the program is scored and a report is generated which highlights the strengths, areas that need improvement, and recommendations for each of the five areas. Program scores are also compared to the average from across all programs that have been assessed.

E. Limitations of the CPC

There are several limitations to the CPC that should be noted. First, the instrument is based upon an “ideal” program. The criteria have been developed from a large body of research and knowledge that combines the best practices from the empirical literature on “what works” in reducing recidivism. Second, as with all applied research, objectivity and reliability are an issue. Although steps are taken to ensure that the information gathered is accurate and reliable, given the nature of the process, decisions about the information and data gathered are invariably made by the assessor(s). Third, the process is time specific. That is, the assessment is based on the program at the time of the assessment. Though changes or modifications may be under development, only those activities and processes that are present at the time of the review are considered for scoring. Fourth, the process does not take into account all “system” issues that can affect the integrity of the program. Lastly, the process does not address why a problem exists within a program.

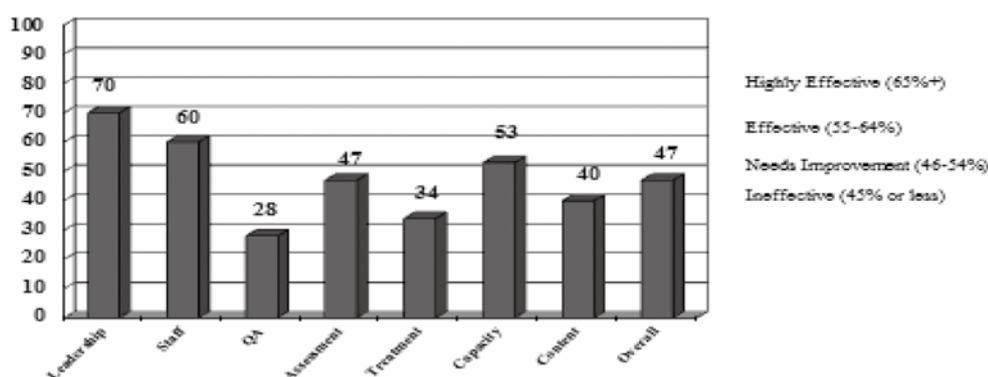
F. Advantages to the CPC

Despite these limitations there are a number of advantages to the CPC. First, it is applicable to a wide range of programs (adult, juvenile, community, institutional, etc.). Second, all of the indicators included in the CPC have been found to be correlated with reductions in recidivism. Third, the process provides a measure of program integrity and quality; it provides insight into the “black box” of a program, something that an outcome study alone does not provide. Fourth, the results can be obtained relatively quickly; usually the process takes a day or two and a report is generated within a few weeks. Fifth, it identifies both the strengths and weaknesses of a program and provides recommendations designed to improve the integrity of the program and to increase effectiveness.

G. Results from the CPAI and CPC

Figure 1 shows the average scores from over 500 program assessments that have been conducted by University of Cincinnati researchers. On average, programs tend to score better in the implementation and staff areas, and poorly in assessment, treatment and quality assurance.

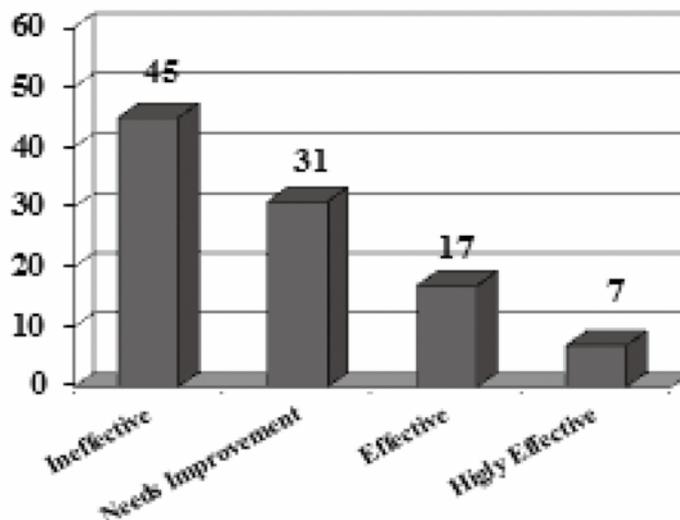
Figure 1. Average Correctional Program Checklist Scores*



*The average scores are based on 512 results across a wide range of programs. Highly Effective = 65% or higher, Effective = 55-64%, Needs Improvement = 46-54%, Ineffective = 45% or less.

Figure 2 shows the overall results from over 500 CPAI and CPC program assessments. The vast majority of programs score in the Ineffective or Needs Improvement categories, while only about a quarter score Effective or Highly Effective. Our research has shown that programs that score in the latter two categories are usually effective in reducing recidivism.

Figure 2. Percentage of Programs in Each Category*



*Based on 512 Assessments across a wide variety of programs. Ineffective=45% or less; Needs Improvement=46-54%; Effective=55-64%; Highly Effective=65% or higher.

While these results indicate that the majority of correctional programs assessed are not fully meeting the principles of effective intervention, they also provide some useful information on how to improve the quality of correctional interventions.

H. Validity of the CPAI and CPC

Several investigators have assessed the predictive validity of the CPAI. Nesovic (2003) reviewed 173 studies (including a total of 266 effect sizes) from the offender treatment literature and reported a mean correlation of $r = .46$ between program scores (i.e. overall CPAI score) and recidivism. Lowenkamp, Latessa and Smith (2006) used the CPAI to conduct 38 reviews of offender treatment programs with matched controls, and reported a mean correlation of $r = .42$ between program scores and recidivism. Lowenkamp also examined the correlations between areas and outcome. Table 2 shows the results. While all areas were significant, program implementation, assessment and treatment were the most important.

Table 2. Correlation of Areas with Recidivism

Area	r (all participants)	r (successful terminations)
Program Implementation	0.56	0.45
Assessment	0.42	0.34
Treatment	0.52	0.39
Staff	0.27	0.33
Evaluation	0.40	0.25
Other	0.16	0.28
Total	0.60	0.47

Over the past several years University of Cincinnati researchers have conducted a number of large outcome studies that examine the relationship between program integrity as measured by the CPAI or the CPC and outcome. These studies included over 40,000 offenders (both adult and juveniles), and over 525 correctional

programs including residential, non-residential, and institutional. Every major study we have done has found a strong relationship between program integrity and recidivism: the higher the integrity scores, the greater the reductions in recidivism. Figures 3 and 4 show the results for both residential and non-residential programs. Residential programs that scored poorly actually increased recidivism by an average of 19 percent, while the highest quality programs reduced recidivism by an average of 22 percent. We see a similar pattern for the non-residential program in Figure 4.

Figure 3. Program Integrity — Relationship between Program Integrity Score and Treatment Effect for Residential Programs

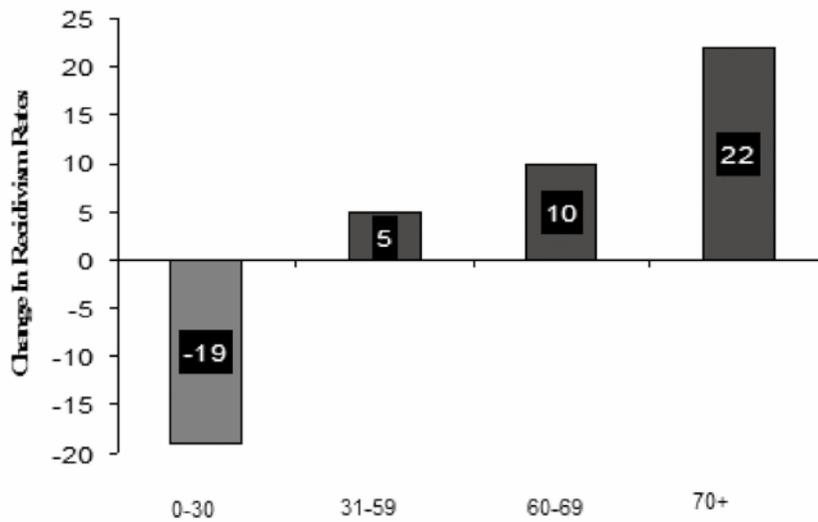
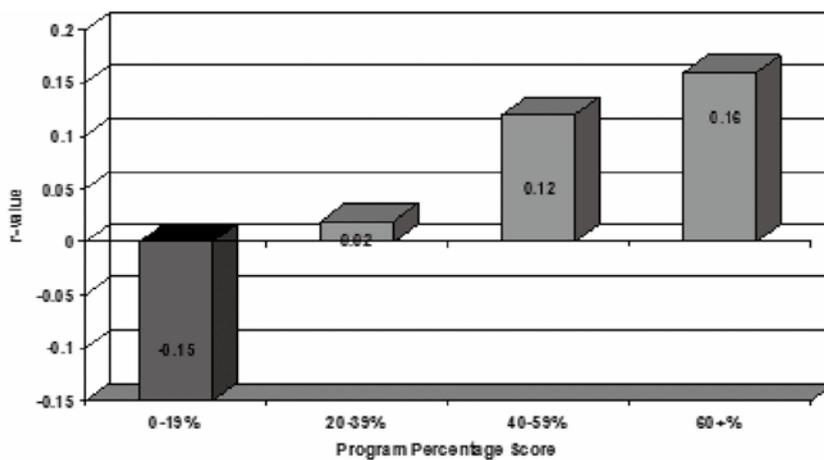


Figure 4. Program Integrity — Relationship between Program Integrity Score and Treatment Effect for Probation/Community Corrections Programs



IV. SUMMARY

Evaluating correctional programs is a challenge. Defining recidivism, measuring outcome, and identifying comparison groups are some of the issues confronting researchers. This paper was meant to provide a brief outline of some of the more common aspects to evaluation research. While evaluation research in corrections is challenging, it can also be rewarding. Building upon our knowledge of what works helps us improve the programs we offer, and enhances our credibility and standing with funding sources and the public.

V. REFERENCES

- Gendreau, P. and D. Andrews (1994). *The Correctional Program Assessment Inventory*.
- Latessa, E. J. and M. M. Moon. 1992. "The Effectiveness of Acupuncture in an Outpatient Drug Treatment Program." *Journal of Contemporary Criminal Justice* 8: 317-331.
- Lowenkamp, C.T. (2004). *Correctional Program Integrity and Treatment Effectiveness: A Multi-site, Program-level Analysis*. Doctoral Dissertation, University of Cincinnati.
- Lowenkamp, C.T., E.J. Latessa and P. Smith (2006). "Does Correctional Program Quality Really Matter? The Impact of Adhering to the Principles of Effective Intervention." *Criminology and Public Policy*, 5 (3): 201-220.
- Lowenkamp, C. T. and E. J. Latessa (2002). *Evaluation of Ohio's Community-Based Correctional Facilities and Halfway House Programs: Final Report*. University of Cincinnati: Center for Criminal Justice Research.
- Lowenkamp, C. T. and E. J. Latessa (2005). *Evaluation of Ohio's RECLAIM Funded Program, Community Correctional Facilities, and DYS Facilities*. University of Cincinnati: Center for Criminal Justice Research.
- Lowenkamp, C. T. and E. J. Latessa (2005). *Evaluation of Ohio's CCA Funded Programs: Final Report*. University of Cincinnati: Center for Criminal Justice Research.
- Nesovick, A. (2003). *Psychometric Evaluation of the Correctional Program Assessment Inventory*. Dissertation Abstracts International, 64 (09), 4674B. (UMI No. AAT NQ83525).
- Osborne, D. and T. Gaebler (1993). *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. New York: Penguin.

APPENDIX A

EVIDENCED BASED CORRECTIONAL PROGRAM CHECKLIST (CPC) SCORING SHEET

Name of Program: _____ Program serves: ___ Males ___ Females ___ Both
 Location (include state): _____ Check program type: ___ Adult ___ Juvenile
 Type of Program: _____ (e.g. institutional, halfway house, day reporting, etc.)
 Primary Treatment: _____ (e.g. substance abuse, sex offenders, general, etc.)
 ___ 1st Assessment ___ 2nd Assessment ___ 3rd Assessment ___ 4th Assessment ___ 5th Assessment
 Date of Assessment: _____ Name of Assessor(s): _____

1. Program Leadership and Development		Check if verified by two or more sources
1.1 PD Qualified	___ 0 or 1	___
1.2 PD Experienced	___ 0 or 1	___
1.3 PD Selects Staff	___ 0 or 1	___
1.4 PD Trains Staff	___ 0 or 1	___
1.5 PD Supervises Staff	___ 0 or 1	___
1.6 PD Conducts Program	___ 0 or 1	___
1.7 Literature Review Conducted	___ 0 or 1	___
1.8 Pilot Interventions	___ 0 or 1	___
1.9 Valued by CJ Community	___ 0 or 1	___
1.10 Value by At-large Community	___ 0 or 1	___
1.11 Funding adequate	___ 0 or 1	___
1.12 Funding stable past 2 years	___ 0 or 1	___
1.13 Program 3 years or older	___ 0 or 1	___
1.14 Gender of groups	___ 0, 1 or N/A	___
SCORE		___ / ___

2. Staff Characteristics		
2.1 Staff Education	___ 0 or 1	___
2.2 Relevant Experience	___ 0 or 1	___
2.3 Staff selected for skills & values	___ 0 or 1	___
2.4 Regular Staff meetings held	___ 0 or 1	___
2.5 Assessed on Service Delivery	___ 0 or 1	___
2.6 Clinical Supervision	___ 0 or 1	___
2.7 Staff Trained on program	___ 0 or 1	___
2.8 On-going Training	___ 0 or 1	___
2.9 Staff input	___ 0 or 1	___
2.10 Staff support treatment goals	___ 0 or 1	___
2.11 Ethical Guidelines for staff	___ 0 or 1	___
SCORE		___ / ___

3. Offender Assessment		
3.1 Appropriate Clients	___ 0 or 1	___
3.2. Exclusionary criteria followed	___ 0 or 1	___
3.3 Risk Factors Assessed	___ 0 or 1	___
3.4. Risk Methods	___ 0 or 1	___
3.5 Risk Level Defined	___ 0 or 1	___

151ST INTERNATIONAL TRAINING COURSE
VISITING EXPERTS' PAPERS

3.6 Need Factors Assessed	___ 0 or 1	___
3.7 Need Methods	___ 0 or 1	___
3.8 Need Level Defined	___ 0 or 1	___
3.9 Responsivity Assessed	___ 0 or 1	___
3.10 Responsivity Methods	___ 0 or 1	___
3.11 Responsivity Defined	___ 0 or 1	___
3.12 Program Targets higher risk	___ 0 or 3	___
3.13 Validation Risk/Needs	___ 0 or 1	___

SCORE ___/___

4. Treatment Characteristics

4.1 Criminogenic targets	___ 0 or 1	___
4.2 Criminogenic target density	___ 0 or 1	___
4.3 Type Treatment	___ 0,1 or 3	___
4.4 Length Treatment	___ 0 or 1	___
4.5 Location monitored	___ 0 or 1	___
4.6 Manual developed	___ 0 or 1	___
4.7 Manual followed	___ 0 or 1	___
4.8 Involvement 40-70%	___ 0 or 1	___
4.9 Groups separated by risk	___ 0 or 1	___
4.10 Intensity varies by Risk	___ 0 or 1	___
4.11 Match Treatment and offender	___ 0 or 1	___
4.12 Match Staff and offender	___ 0 or 1	___
4.13 Match Staff and program	___ 0 or 1	___
4.14 Offender Input	___ 0 or 1	___
4.15 Use Appropriate Rewards	___ 0 or 1	___
4.16 Ratio Favors Rewards	___ 0 or 1	___
4.17 Procedures for rewards	___ 0 or 1	___
4.18 Appropriate punisher	___ 0 or 1	___
4.19 Procedure for Punishment	___ 0 or 1	___
4.20 Negative Effects	___ 0 or 1	___
4.21 Completion Criteria	___ 0 or 2	___
4.22 Completion rate	___ 0 or 1	___
4.23 Skills Modeled	___ 0 or 1	___
4.24 Skill training	___ 0 or 1	___
4.25 Graduated practice	___ 0 or 1	___
4.26 Groups monitored by staff	___ 0 or 1	___
4.27 Group size	___ 0 or 1	___
4.28 Significant Others trained	___ 0 or 1	___
4.29 Discharge planning	___ 0 or 1	___
4.30 Aftercare provided	___ 0 or 1	___
4.31 Quality aftercare	___ 0 or 1	___

SCORE ___/___

5.0 Quality Assurance

5.1 Internal Quality Assurance	___ 0 or 1	___
5.2 External Quality Assurance	___ 0 or 1	___
5.3 Client Satisfaction	___ 0 or 1	___
5.4 Offenders reassessed	___ 0 or 2	___

RESOURCE MATERIAL SERIES No.88

5.5 Recidivism tracked _____ 0 or 1 _____
 5.6 Program evaluated _____ 0 or 1 _____
 5.7 Program effective _____ 0 or 1 _____
 5.8 Evaluator working with program _____ 0 or 1 _____

SCORE _____/_____

TOTAL SCORE _____/_____

CAPACITY AREAS: Leadership & Development _____ %
 Staff _____ %
 Quality Assurance _____ %
 CONTENT AREAS: Assessment _____ %
 Treatment _____ %
OVERALL CAPACITY _____ %
OVERALL CONTENT _____ %
OVERALL _____ %

OVERALL RATING: _____
 1= Highly Effective (65% or higher)
 2= Effective (55-64)
 3= Needs Improvement (46-54%)
 4= Ineffective (45% or higher)