The Importance of Evaluating Correctional Programs: Assessing Outcome and Quality

Edward J. Latessa and Alexander Holsinger

The importance of evaluating correctional programs, especially those operating in the community, has never been more pronounced. Vast sums of money are spent on correctional programming, and the public is demanding programs that work. The critical question of course is indeed "what works"? This article examines some of the issues surrounding program effectiveness: what is important to measure, what do we know about program effectiveness, and what harm is done when we fail to develop effective programs.

Key words: *correctional intervention, evaluation, principles of effective intervention, program quality, recidivism*

Corrections Management Quarterly, 1998, 2(4), 22–29
© 1998 Aspen Publishers, Inc.

One of the most important areas of contemporary concern for correctional officials is the design and operation of effective correctional intervention programs. This is particularly relevant since there is consistent evidence that the public supports rehabilitation programs for offenders.[1,2] Furthermore, survey research also reveals strong support for public protection as an important goal of corrections.[1] Consequently, disagreements arise, focusing on what are the best methods to achieve these and other correctional goals. On one side are advocates for more punitive policies such as an increased use of incarceration, punishing smarter (e.g. boot camps, electronic monitoring, day reporting facilities), or simply increasing control and monitoring if the offender is supervised in the community. Those advocating these strategies of crime control do so on the basis of the often interrelated goals of punishment—deterrence, retribution, and incapacitation. The limits of these approaches have been outlined and debated by others.[2,3] The ultimate effect of most of these policies has been described as "penal harm" by Clear.[4]

As Cullen and Applegate imply, the most disheartening aspect of these "get tough" policies is their dismissal of the importance of programming that is designed to rehabilitate offenders.[5] Cullen and Applegate further question whether the penal harm movement's rejection of rehabilitation is sound public policy. As many states have found, simply locking up offenders and "throwing away the key" has proven to be a very expensive approach to crime con-

Edward J. Latessa, PhD, is Professor and Head, Division of Criminal Justice, University of Cincinnati, Cincinnati, Ohio.

Alexander Holsinger, MS, is Research Associate, Division of Criminal Justice, University of Cincinnati, Cincinnati, Ohio.

trol. This approach is also very limited, since the vast majority of offenders will one day return to society. Many will return unchanged at best, and, at worst, with many more problems and intensified needs for service.[6] For those advocating incapacitation, one must also ask what should be done with offenders while incarcerated? Some scholars, such as Cullen and Applegate, do not believe that incapacitation and rehabilitation are mutually exclusive. Furthermore, since the vast majority of offenders are supervised in the community at differing degrees of intensity, it is even more important that we develop programs that work toward reducing recidivism.

Many of the "intermediate sanctions" that have been developed over the past few years are but a few examples of "programs" that often fail to live up to their expectations, particularly in terms of reductions in recidivism.[7,8] These results are often attributed to policies that emphasize control and surveillance over treatment and service delivery.[9,b]

Despite the punitive movement, there is increasing evidence that correctional treatment can be effective in reducing recidivism among offenders.[10–13] Nonetheless, some scholars remain unconvinced.[14–16,c] The debate surrounding treatment effectiveness has been on going since Martinson's proclamation that "nothing works," with many still clinging to this mantel despite evidence to the contrary. Principal among the reasons for disbelief in the potential effectiveness of correctional programming are the failure to measure outcome properly and the lack of quality programs.

Measuring Outcome and the Limitations of Recidivism

Correctional outcome, which is usually operationalized as recidivism, has some inherent limitations. The indicators used to measure recidivism, length of follow-up, and external and internal factors affect recidivism rates. Indeed, the best way to ensure a low recidivism rate is to define it very narrowly (e.g., incarceration in a state penal institution) and to utilize a very short follow-up period.

Too often, arrest (and only arrest) is used as a primary indicator when measuring recidivism and, consequently, program outcome. Certainly, arrest may serve as an indicator of postprogram (or postrelease) performance, but in and of itself arrest has many limitations. Some of the other factors that are overlooked when considering the impact of a correctional program or criminal sanction, even when arrest is being used, are time until arrest; offense for which an offender was arrested (type of offense as well as severity level); whether or not the offender was convicted; and, if convicted, what the resulting disposition was.

Despite these limitations, recidivism remains the most important measure of public protection. When legislators and other public officials ask if a program works, they are generally referring to recidivism. Outcome studies provide much of our knowledge about the effectiveness of correctional programs in reducing recidivism. Unfortunately, outcome studies focus on the results of intervention and provide little, if any, useful information about why a program is or is not effective. In addition to the measurement of outcome, another factor that can influence recidivism rates is the quality of a program.

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 has been the development of criteria or indicators by which a treatment 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 if 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 assessment 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.

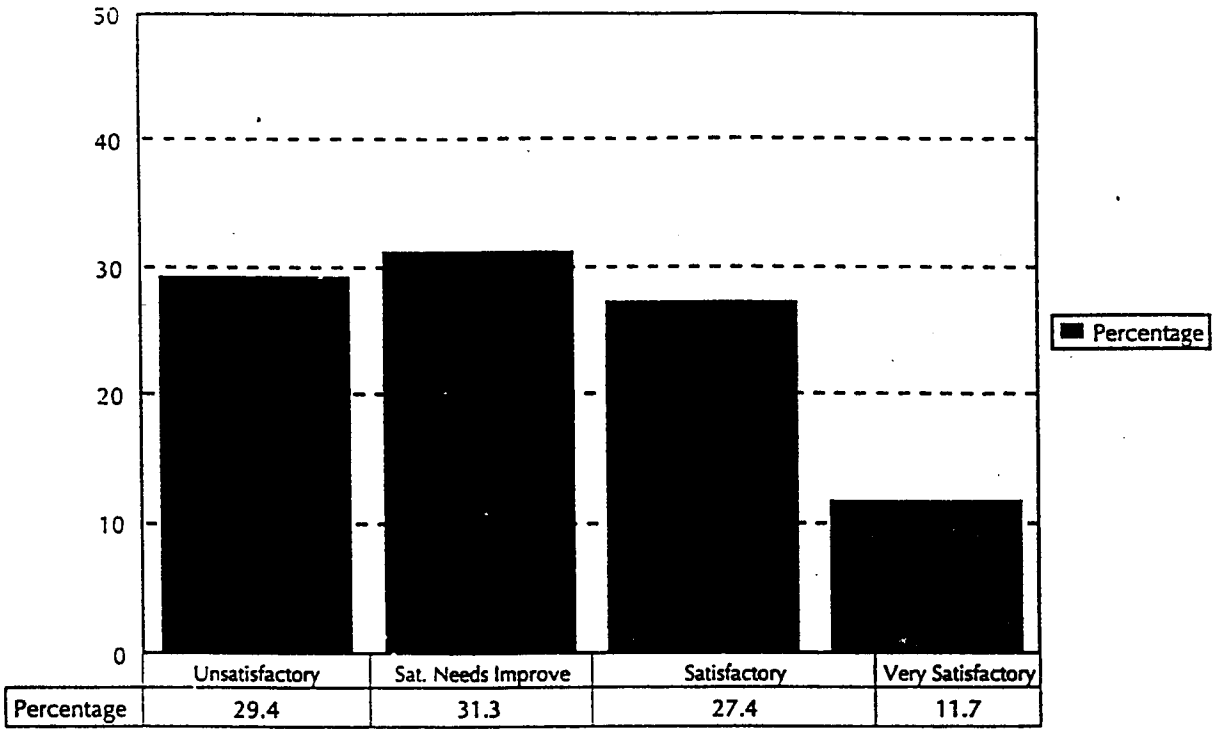

Figure 1. Average Correctional Program Assessment Inventory scores. The average scores are based on 51 CPAI results across a wide range of programs. The CPAI was developed and is copyrighted by Paul Gendreau and Don Andrews.

Principles of Effective Intervention

As mentioned above, evidence shows that many rehabilitation programs have in fact produced significant reductions in recidivism. The next critical issue became the identification of those characteristics most commonly associated with effective programs. Through the work of scholars such as Andrews and colleagues[10] and Gendreau,[11] several "principles of effective intervention" have been identified. The following statements summarize these principles of effective intervention:

- Services should be intensive, behavioral in nature, and focused on higher risk offenders.
- Behavioral strategies should be enforced in a firm but fair manner by qualified staff.
- Programs should target criminogenic needs (e.g., antisocial attitudes, value and beliefs supportive of criminal behavior, negative peer associations, substance abuse, etc.).
- Responsivity should occur between staff, offenders, and programs.
- Programs should disrupt criminal networks.
- Programs should provide relapse prevention in the community.
- High levels of advocacy and brokerage should be provided.

In order to determine whether or not a program is meeting these principles, each of them need to be operationalized into a set of standards or guidelines. Gendreau and Andrews[17] accomplished this through the development of an instrument known as the Correctional Program Assessment Inventory (CPAI).

The CPAI is a tool that has been used for assessing correctional intervention programs. The CPAI assesses a program on six primary areas: 1) program implementation and leadership; 2) offender assessment and classification; 3) characteristics of the program; 4) characteristics and practices of the staff; 5) evaluation and quality control; and 6) miscellaneous items such as ethical guidelines and levels of support. Each section is scored as either "very satisfactory" (70 percent to 100 percent), "satisfactory" (60 percent to 69 percent), "satisfactory but needs improvement"

(50 percent to 59 percent; or "unsatisfactory" (less than 50 percent).[d] The scores from all six areas are totaled and the same rating scale is used for the overall assessment score. Not all of the six areas are given equal weight, and some items may be considered "not applicable," in which case they are not included in the scoring.

To date, researchers from the University of Cincinnati have assessed more than 50 correctional programs across the country. The average scores from across 51 programs are presented in Figure 1.[e] In two of the areas, including the most important (treatment), the average scores were less than 50 percent, which is in the unsatisfactory range of the scale. The assessment and staff scores were slightly higher, falling in the "satisfactory but needs improvement" range. Implementation and the "Other" category were found to be "very satisfactory." Overall, the 51 programs we assessed averaged just fewer than 56 percent on the CPAI. Figure 2 shows the percentages of programs that scored in each category. Only 11.7 percent of the programs we assessed scored in the "very satisfactory" range, while 60 percent scored either "satisfactory but needs improvement" or "unsatisfactory."

Although there is a great deal of variation between (and within) programs, we have summarized some of the major strengths and weaknesses in each of the CPAI areas across the 51 programs we have assessed.

Program implementation and leadership

Strengths: Effective programs have strong leadership and involvement of the program director.[f] For the most part we found qualified and experienced program directors, who were involved in designing the program. They tend to be involved in the hiring and training of staff, and in many instances they provide some direct services to offenders. It is also important for the survival and support of a program that the values and goals of the program be consistent with existing values in the community or institution in which it resides and that there be a documented need for the program. Support for the program also depends on perceptions of cost-effectiveness. We usually find that most correctional programs meet these conditions.

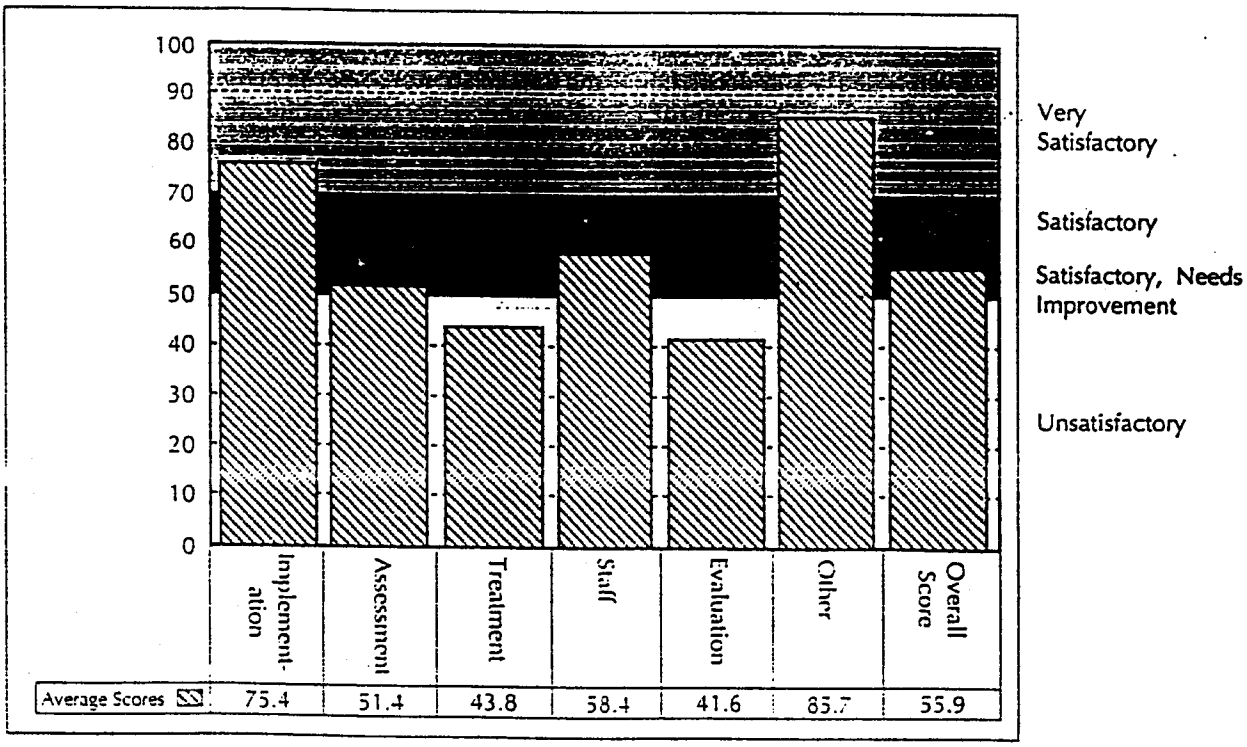

Figure 2. Programs in each CPAI category (based on 51 CPAI assessments across a wide variety of programs.

Weaknesses: There are two flaws that are common in this area. Effective programs are based on strong theoretical models that are derived from the treatment literature. Regardless, many of the correctional intervention programs we examined were basically designed with little regard for the empirical research on what works with the type of offenders they were serving. In addition, effective programs are usually started on a pilot basis to work out the logistics. Thus far, we have found few programs that piloted their treatment components before full implementation.

Offender assessment and classification

Strengths: The vast majority of programs we studied have stated criteria for admissions, receive appropriate clients, and have a rational legal/clinical basis for the exclusion of certain types of offenders. We also found that, in general, most programs attempt to assess some offender characteristics related to risk and need.

Weaknesses: While many programs did indeed attempt to assess offenders regarding risk and need, doing so did not involve the incorporation of a standardized, objective, actuarial instrument. The absence of actuarial risk/needs assessment instruments was particularly evident in programs that deal with juvenile offenders. Even when a standardized assessment is being performed at some point in the offender's entry/progress, it is seldom found that the information gathered is being used to distinguish offenders by risk. In other words, even when proper (and potentially beneficial) assessments are being performed, the information is not influencing the decision-making process, let alone service delivery. In addition, it is generally found that staff assessments of offenders are based on a quasiclinical approach that does not result in a summary score. Likewise, it has been very rare to find that programs are routinely measuring with standardized instruments responsivity characteristics, such as levels of motivation, intelligence, or psychological development.

Characteristics of the program

Strengths: This is a critical area in scoring the CPAI. Unfortunately it is also one that typically results in low scores.[3] Effective intervention programs focus the vast majority of their efforts on targeting criminogenic needs and behaviors.[h] In general we have found that many correctional intervention programs target these behaviors (although we still find programs that provide intensive services and treatment in non–crime producing areas, such as self-esteem). Another common strength was that many programs have criteria for program completion, and upon discharge many offenders are routinely referred to programs and services that help meet their needs.

Weaknesses: Offenders typically have not been spending a significant percentage of their time in structured programs. In addition, the amount of services and treatment provided has not been varying by risk and need levels. Yet another characteristic of an effective program is the use of a treatment model that has been found to be effective. Since programs are rarely designed around a theoretical model, it was not surprising to find a lack of a consistently applied treatment model in place. In general, the major shortcomings found when considering the "Characteristics of Program" portion of the CPAI include lack of programmatic structure; incomplete or nonexistent treatment manuals; few rewards to encourage program participation and compliance; the ineffective use of punishment; staff being allowed to design their own interventions regardless of the treatment literature base; and a host of very obvious and definable, yet ineffective, treatment models. This area of the CPAI also examines the extent to which matching occurs between offenders and staff, offenders and programs, and staff and programs. Even when matching is found to occur, it is uncommon to observe it being based on specific responsivity criteria. In addition, it is very rare to find a program that includes family and/or friends of the offender in the treatment process. Finally, many programs failed to provide aftercare services or booster sessions.

Characteristics and practices of the staff

Strengths: Although there is a great deal of variation from program to program regarding staff quality, for the most part we found educated and experienced staff working with offenders. Often staff were selected on personal characteristics such as life experience, fairness, firmness, and problem-solving skills. We also found that staff usually had input in the structure of the programs and that on-going training was provided.

Weaknesses: Staff turnover was often a problem with some types of correctional programs (e.g., halfway houses), and we rarely found staff who had re-

ceived sufficient training on the interventions and treatments utilized by the program. Clinical supervision was not routinely provided, and staff were rarely assessed on service delivery skills.

Evaluation and quality control

Programs that study themselves tend to be more effective than programs that do not. Data provides insight into program and offender performance, helps identify who is successful and who is not, and allows adjustments to be made.

Strengths: File review and case audits were usually conducted.

Weaknesses: Periodic, objective, and standardized assessment of offenders to see if criminogenic factors were being reduced was uncommon. In short, most programs do not develop meaningful performance measures (to measure either program or offender performance over time). We also found that the majority of programs were not tracking offenders after they have left the program, and formal evaluations involving comparison groups were the exception.

Other items

Strengths: Most of the programs we examined score well in this area. In general, offender records are complete and are kept in a confidential file. Changes that jeopardize programs, funding, or community support are rare.

Weaknesses: Some programs do not have ethical guidelines for intervention; public agencies tend not to have advisory boards, while those operated by nonprofits do.

While these results indicate that the majority of correctional programs we 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. The most common shortcomings of programs we have assessed include the following:

- Programs tend to be atheoretical.
- A lack of standardized, objective risk/need assessments (especially with juveniles) are conducted, and everyone gets the same treatment.
- Little attention is given to responsivity between staff, offenders, and programs.
- Staff training is inadequate, and the quality of staff varies greatly.
- There are too few rewards and too many punishers. However, the punishers are not effective because of how they are administered.
- Few measures of program performance are taken.
- Family members are not involved in the treatment process.
- Aftercare and booster sessions are lacking.
- Few formal evaluations are conducted.

There are several advantages to evaluating a correctional treatment program using the CPAI. First, the CPAI is a versatile assessment tool applicable to a wide range of programs (adult, juvenile, community, institutional, etc.). Second, the instrument allows for the quality of a program to be quantified through a scoring process. This allows for comparisons between programs, as well as benchmarking. Third, the instrument is relatively efficient and can be completed in a short period of time. Finally, the CPAI can be used as an evaluation tool. Programs that do not achieve at least a "satisfactory" score on the instrument would probably not be expected to reduce recidivism. Investment in an outcome study would be inadvisable until the program made the changes necessary to reach an acceptable level of program quality.

Treatment Harm

The previous section outlined many of the pertinent issues of evaluating program effectiveness. It has become apparent that the way recidivism is measured may need to be reconsidered in order to more fully understand the essence of a program's crime-reducing potential. Further, and perhaps more importantly, the CPAI allows for measures of program quality, across many different areas, to be included in future program evaluations. However, given that not all areas of the country are yet aware of the CPAI, it is quite likely that ineffective programs will continue to receive new and/or continued financial support. The result of allowing ineffective correctional strategies to continue is quite likely to be treatment harm, which manifests itself in many ways.

- *The undermining of confidence and support.* When faced with the existence of ineffective correctional treatment strategies, judges, legislators, and the public lose confidence in treatment efforts, which in turn undermines support for effective programs. In the realm of public opin-

ion regarding correctional treatment, some bad apples may indeed spoil the entire bushel—programs that are doing good treatment based on what we know works can be harmed by programs that are not. In short, if ineffective treatment strategies continue to flourish, the risk of returning to a "nothing works" doctrine will undoubtedly increase.

- Empowers fartcatchers. Gendreau[18] has used this term to describe those staff members, bureaucrats, and politicians who lack substantive knowledge in criminal behavior and criminal justice. These "MBA types" are often making policy on the basis of limited or "quick fix" solutions that have empirical support. The result has often been programs that promise a quick fix to the problem of criminal behavior, or what Cullen refers to as correctional "quackery." There is no panacea approach to changing offender behavior. Those searching for a "magic pill" for reducing crime and changing human behavior will not only be disappointed but will, at best, perpetuate the myth that there is some simple solution and, at worst, erode faith in the potential for effective intervention.
- Offenders get all the blame. Perhaps the greatest harm that is perpetuated is what is done to the offender. When we use programs that are not effective, the offender receives the blame, not the program. Too often we hear judges say that they have already sent the offender to some program and that they have failed. This "failure" is then used to justify even harsher punishments. Sending an offender to a program that does not adequately meet the principles of effective intervention should not be expected to change criminal behavior. Blaming the offender absolves us from designing and operating high-quality and effective intervention programs.

What Can Be Done To Evaluate Correctional Intervention Programs?

First, measure outcome. Recidivism cannot be ignored. Use multiple indicators (arrests, convictions, incarceration rates, program completion, type and severity of offenses, and technical violations). Conduct follow-up studies of sufficient length to adequately assess recidivism. Develop comparison groups. While experimental designs are the most powerful approaches, do not hesitate to use quasi-experimental research designs if necessary to develop comparison groups.

Second, measure program quality. Get inside the "black box" of an intervention program and do not hesitate to ask about program quality. Are the programs you are operating or funding of sufficient quality to reduce the criminal behavior of the offenders it is serving? If so, what evidence exists to prove this?

Third, make evaluation an integral part of your program, not simply an afterthought. A good evaluation can assist your program in many ways. You should also form partnerships with researchers. Developing relationships with evaluators and researchers can yield valuable returns that often go beyond research reports. Evaluators can help to develop databases, provide interns, assist in developing offender selection processes and criteria, help in the identification and selection of assessment and classification instruments, and ensure that important information about the offenders being served and the program itself are being captured in a useful form.

Fourth, do not ignore established programs. Often it is the newer programs that undergo evaluation scrutiny. Rarely are older programs subjected to evaluation. This is a mistake for several reasons. First, all programs can be improved, and data from evaluations can provide useful information for program change. Second, just because a program has been in operation for a long time does not mean that it is effective. Again, are we perpetuating correctional "quackery" by failing to assess and evaluate programs?

Finally, use data to improve your program. Unfortunately, reports that collect dust on shelves are often the rule rather than the exception. Programs that are data driven tend to be more effective than those which are not.

• • •

What do we know from the correctional research? When we review the body of literature on program effectiveness, there are several important points that can be made. First, programs that meet the principles of effective intervention are more effective than programs that do not. Second, the importance of sound empirical research and periodic evaluation of correc-

tional programs cannot be stressed enough. Without empirical evidence we will continue to operate and support programs that do not work. Our failures cancel out our successes,[13] and we will be unable to distinguish between our successful programs and our failures. Finally, as part of our evaluation efforts we cannot afford to ignore the measurement of program integrity.

ENDNOTES

a. This study found results similar to other studies and concluded that support for rehabilitation, which appeared to have weakened in the 1970s, may have stabilized in the intervening 15 years.
b. Those programs that have shown some reductions in recidivism have adopted a more balanced approach to offender supervision. With this model, an equal emphasis is placed on monitoring the offender and delivering services.
c. Lab and Whitehead concluded that there was not evidence that treatment programs for juvenile offenders were not effective despite the fact that 40 percent of the programs they had reviewed found reductions in recidivism. For an excellent rebuttal to Logan and Gaes, see Cullen and Applegate (1998) "Introduction" in *Offender Rehabilitation: Effective Correctional Intervention.* Brookfield, VT: Ashgate Dartmouth: xviii–xix.
d. The "satisfactory but needs improvement" category was created because of what we felt were substantial differences between a program that scored in the sixties and one that scored in the fifties. In actuality, all scores in these ranges need improvement.
e. The programs include adult and juvenile, and community based as well as institutional.
f. Program director refers to the individual responsible for service or treatment delivery. This is not necessarily the administrator of the program.
g. The Program Characteristics area of the CPAI accounts for about one-third of the total score.
h. Criminogenic needs refer to those areas that are found to be strong predictors of criminal behavior, such as antisocial attitudes, criminal values, and beliefs, negative peer associations, etc.

REFERENCES

1. Applegate, B. K., Cullen, F.T., and Fisher, B.S. "Public Support for Correctional Treatment: The Continuing Appeal of the Rehabilitative Ideal." *The Prison Journal* 77 (1997): 237–258.
2. Currie, E. *Confronting Crime: An American Dilemma.* New York: Pantheon, 1985.
3. Bennett, W. J., Dilulio, Jr., J.J., and Walters, J.P. *Body Count: Moral Poverty...and How to Win America's War Against Crime and Drugs.* New York: Simon and Schuster, 1996.
4. Clear, T.R. *Harm in American Penology: Offenders, Victims, and Their Communities.* Albany, NY: Albany State University of New York Press, 1994.
5. Cullen, F. T., and Applegate, B.K. *Offender Rehabilitation: Effective Correctional Intervention.* Brookfield, VT: Ashgate Dartmouth, 1998.
6. Petersilia, J. "California's Prison Policy: Causes, Costs, and Consequences." *The Prison Journal* 72 (1992): 8–36.
7. Latessa, E. J., Travis, L.F., and Holsinger, A. *Evaluation of Ohio's Community Correction Act and Community-Based Correctional Facilities.* Cincinnati, OH: Division of Criminal Justice, University of Cincinnati, 1997.
8. Petersilia, J. "Probation in the United States." In M. Tonry, ed. *Crime and Justice: A Review of Research,* Vol. 22 (pp. 149–200). Chicago: University of Chicago Press, 1997.
9. Fulton, B., Latessa, E.J., Stichman, A., and Travis, L.F. "The State of ISP: Research and Policy Implications. *Federal Probation* 61, no. 4 (1997): 65–75.
10. Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., and Cullen, F.T. "Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis." *Criminology* 28 (1990): 97–104.
11. Gendreau, P., and Andrews, D. "Tertiary Prevention: What the Meta-Analysis of the Offender Treatment Literature Tells us About 'What Works.' *Canadian Journal of Criminology* 32 (1990): 173–184.
12. Cullen, F. T., and Gendreau, P. "The Effectiveness of Correctional Rehabilitation: Reconsidering the 'Nothing Works' Debate." In L. Goodstein and D. MacKenzie, eds. *American Prisons: Issues in Research and Policy* (pp. 23–44) New York: Plenum, 1989.
13. Van Voorhis, P. "Correctional Effectiveness: The High Cost of Ignoring Success." *Federal Probation* 51, no. 1 (1987): 56–62.
14. Antonowicz, D. H., and Ross, R.R. "Essential Components of Successful Rehabilitation Programs for Offenders." International *Journal of Offender Therapy and Comparative Criminology* 38 (1994): 97–104.
15. Lab, S., and Whitehead, J. "An Analysis of Juvenile Correctional Treatment." *Crime and Delinquency* 28 (1988): 60–85.
16. Logan, C. H., and Gaes, G.G. "Meta-Analysis and the Rehabilitation of Punishment." *Justice Quarterly* 10 (1993): 245–263.
17. Gendreau, P., and Andrews, D. *The Correctional Program Assessment Inventory* (5th ed.). Saint John: University of New Brunswick, 1994.
18. Gendreau, P. Plenary Speech to the American Probation and Parole Association, Boston, Massachusetts, 1997.