

Applied research in support of program effectiveness

Franca Cortoni¹

Research Branch, Correctional Service of Canada

Conducting applied research to determine the effectiveness of programs in addressing criminogenic factors and reducing reoffending is too often viewed and treated as a process that takes place well after a program has been in operation and sufficient time has elapsed to determine outcome. This situation, however, leads to very unsatisfactory results as during the evaluation questions inevitably arise that cannot be answered. The design of applied research that aims to assess the effectiveness of programs must take place at the developmental stages of the program, and be reviewed and adjusted as changes to the program takes place.² This process ensures that safeguards are in place to prevent errors and to allow for valid conclusions to be drawn.³

The Logic Model

The general goal of an evaluation is to measure change.⁴ The research design is established to verify the outcome, and verify that this outcome (i.e., the change) is due to the program. A solid research design ensures that the expected changes are carefully defined and operationalized. Further, it ensures that other variables are controlled so that the change can be clearly attributed to the program.

The Logic Model is a useful framework to help establish the link between objectives and expected outcomes of the program.⁵ It also provides a good source of information to inform what other variables, such as length of incarceration and other program participation, may be related to the outcome.⁶

Controlling for these variables is crucial to demonstrate that the outcome is due to the specific program and not to some other reason. A Logic Model is a flow diagram that establishes the links among the main components of the program, the implementation objectives, and the short and long outcome objectives. It allows for planned decision-oriented evaluations. Research questions and methods of investigation flow naturally from the model and activities related to the research are clarified. Some of these activities include identifying who will use the research results and for what purposes; estimating the required short and long term resources to conduct the research; and establishing methods for data collection. The model also serves as the basis for identifying intermediate and long-term outcome evaluations. Finally, it helps

ensure that all aspects related to the program design, implementation, and evaluation are clarified a priori. In other words, using a Logic Model as the initial step in designing an evaluation framework promotes the clarification of objectives; identifies missing components; distinguishes between means and ends; ensures that the intended outcomes are clearly related to the program components and are clearly established; and removes assumptions about what the program is to achieve.⁷

Defining the research questions

The standard question in evaluating correctional programs is whether the program was successful in reducing reoffending. To answer that question, well-designed evaluations with long-term follow-ups are required. Despite this, it could be difficult to establish such effectiveness on the basis of single studies, particularly those that evaluate reductions in sub-types of offending behaviour that occur in low frequency such as general and sexual violence. It must be remembered, however, that a single carefully designed outcome evaluation not only informs about the current program, but also adds to the wider body of research and may become part of later meta-analytical studies. Meta-analytical research is rapidly becoming the method of choice to demonstrate that programs are effective in reducing reoffending. These studies also help reduce the impact of threats to internal and external validity found in non-randomized evaluative designs.⁸

Whether a program impacted positively on successful risk management is not, however, the only question of relevance. Applied research also examines whether the program was successful in enrolling the appropriate candidates; was effective in addressing its treatment targets; and whether immediate or short-term changes, such as improved institutional adjustment, were achieved. On the longer term, the maintenance of treatment gains needs to be verified, as well as the relationship between these treatment gains and increased prosocial behavior in the community. Finally, it is useful to simply ask participants what they thought of the program and whether they felt it addressed their need.

Psychometric instruments

Psychometric instruments, or paper-and-pencil tests, are routinely utilized in program evaluations. These instruments assess offenders on the aspects that are expected to change as a result of program participation. The tests are typically administered before the start of treatment, and following completion of the program. If the program was successful, there should be differences between the results of the pre- and the post-program tests. For example, if the program is successful in reducing anti-social attitudes, the post-treatment scores on a psychometric instrument that assesses the extent of anti-social attitudes should be lower than the pre-treatment scores. It is those differences that provide one type of evidence that the program is achieving its intended changes.

The importance of psychometric instruments in measuring treatment gains and providing evidence of the program effectiveness is often underestimated. Determining what needs to be measured is also not always carefully considered. Instruments are sometimes selected on the basis of pre-conceived notions about what 'should change' as a result of treatment instead of what is expected to change. For example, too often the assessment of assertiveness is routinely conducted as part of a program, and yet there is no evidence that a lack of assertiveness is related to criminal behaviour. Psychometric instruments should be carefully selected for their reliability and validity in measuring the construct of interest. And the construct of interest should be the criminogenic factor, or an aspect thereof, that is addressed by the program.

Sometimes, no appropriate psychometric test exists for the purpose at hand. In such cases, the tendency is to either include instruments that inadequately measure the construct, for example a personality test, or to create a new instrument. Research shows that the use of inadequate instruments does not provide meaningful information regarding the effectiveness of the program.⁹ The creation of a new instrument may be very appropriate but requires several stages of development to establish its reliability and validity.¹⁰ It is this empirical development that is too often neglected when new tests are being developed. If the psychometric properties of a test are not established, any subsequent research using the results of the non-validated test becomes meaningless. Hence, the verification of the new test's ability to appropriately measure the construct must be part of the research framework.

The problem of comparison groups

Untreated comparison groups should be as closely matched as possible to the treatment group. Relevant matching variables involve at a minimum: age, overall risk level; and identified criminogenic needs. Issues such as settings, time and length of sentence should also be considered. Setting refers to matching on the conditions surrounding the treated and the comparison offenders such as the type of institution where the program took place. Setting is a variable mostly ignored in applied research.¹¹ Yet its impact could be significant if, for example, one is interested in examining the outcome of programs that aim, as an intermediate goal, to improve institutional adjustment, or whether a residential treatment program is equally effective if delivered in a regular penal institution.

Likely the single most difficult issue in terms of creating comparison groups is controlling for participation in programs other than the one under study. In the Correctional Service of Canada (CSC), offenders generally complete a range of correctional programs, all designed to address various criminogenic factors. For example, federally sentenced sexual offenders typically take an average of 3.2 different types of programs.¹² It is only in rare cases that offenders do not take any program, either institutionally or in the community. It is clear then that any evaluation of the impact of a specific program necessitates that participation in other programs be taken into account. As Los el notes, the more therapeutic the prison environment, the more difficult it is to establish that a particular program is effective.¹³ This situation attests to the importance of creating appropriate comparison groups, and using innovative methods to control for the effect of participation in other therapeutic or correctional interventions.

Matching offenders on program participation, in addition to other relevant variables, is extremely problematic. Offenders would not only need to be matched on type and intensity of programs, but also on whether they were successful or not in completing these programs. Innovative strategies to control for these variables are therefore required. Dowden and Serin developed one such strategy in their evaluation of the Anger Management Program.¹⁴ They created a composite program performance variable that took into account both the number of programs undertaken by offenders and whether they were successful or not in completing these programs. They then matched offenders in the treatment and the comparison group on this variable to control its effect. This technique permitted the authors to conclude that the participation in the

Anger program contributed uniquely to the subsequent lower rates of violent and general reoffending. As this example shows, when the comparison group is well matched, more robust conclusions may be drawn from the research.

The issue of treatment attrition

A program may be highly effective in reducing recidivism for those offenders who complete it. However, the cost-effectiveness of the program may be compromised if only a small proportion of offenders complete it.¹⁵ Further, studies show that offenders who start but fail to complete treatment reoffend at a higher rate than offenders who either complete treatment, or do not start at all.¹⁶ Program attrition is a universal phenomenon that is well recognized in the correctional literature. In CSC, while there is variety in actual rates, an average of 20% of offenders who enter programs fail to complete for a number of reasons,¹⁷ a statistic that compares favorably with dropouts rates reported in the literature.¹⁸

Typically, offenders drop out of programs because their expectations were not met. They found the program too demanding, or they did not believe the program would help them. In addition, program participants may be removed from a program for being disruptive or due to lack of attendance. Taken together, the reasons for dropping out of treatment are often seen as indicative of a lack of motivation on the part of the offender. A lack of motivation may be related to resistance in the client,¹⁹ or related to a failure to envision the intrinsic benefits of

participating in the program.²⁰ The evaluative design should therefore include methods to assess motivation and determine its impact on program participation.

Another potential cause of attrition is a lack of correspondence between the referral and the program. For example, if an offender is assessed as having a lower need in a given area, and yet is required to complete a program of a higher intensity, he may become discouraged, and fail to complete the program. Here again, the Logic Model would specify the targeted population and the system requirements that would ensure only appropriate candidates would be included in the program. The evaluation of the program would then verify that the program was offered to the appropriate candidates. If the wrong offenders receive the program, the subsequent outcome evaluations may be rendered invalid.

Conclusion

Designing and implementing appropriate research to evaluate correctional programs can be a daunting task. Yet, it is the only way to verify that the intervention achieved its stated goals. The overall goal of correctional program intervention is to address criminogenic factors in efforts to reduce rates of reoffending and contribute to the protection of society. Given the importance of this goal, it is incumbent upon program evaluators to develop appropriate research designs that will contribute to our knowledge in this area. ■

¹ 340 Laurier Avenue West, Ottawa, Ontario K1A 0P9

² Patton, M.Q. (1997). *Utilization-focused Evaluation*. Thousand Oaks, CA: Sage Publications Inc.

³ McGuire, J. (2000). Development of a program Logic Model to assist evaluation. In Motiuk, L.L., & Serin, R.C. (Eds.), *Compendium 2000 on Effective Corrections*, Vol. I (pp. 208–220). Ottawa, ON: Correctional Service of Canada.

⁴ Ibid.

⁵ Ogborne, A.C. (2003). *Program Evaluation: An overview of some basic issues*. Paper presented at the Correctional Service of Canada Addiction Research Centre Summer Institute, Montague, Prince Edward Island, Canada.

⁶ Op.Cit., McGuire (2000).

⁷ Op.Cit., Ogborne (2003).

⁸ Los el, F. (2001). Evaluating the effectiveness of correctional programs: Bridging the gap between research and practice. In Bernfeld, G.A., Farrington, D.P., & Leschied, A.W. (Eds.), *Offender Rehabilitation in Practice* (67–92). New York, NY: John Wiley & Sons.

⁹ Ibid.

¹⁰ Nunally, J.C. (1967). *Psychometric Theory*. New York, NY: McGraw-Hill Book Company.

¹¹ Op. Cit., Los el (2001).

¹² Mailloux, D., & Serin, R. (2003). *Issues impacting on programming for sex offenders*. Unpublished paper. Ottawa, ON: Correctional Service of Canada.

¹³ Op. Cit., Los el (2001).

¹⁴ Dowden, C. & Serin, R. (2001). *Anger management Programming for Offenders: The impact of program performance measures*. Research Report R-106. Ottawa, ON: Correctional Service of Canada.

¹⁵ Marshall, W.L., & Williams, S. (2000). Assessment and treatment of sexual offenders. *Forum on Corrections Research*, 12(2), 41–44.

¹⁶ Hanson, R.K., & Bussiere, M.T. (1996). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 63, 802–809. Also see, Robinson, D. (1995). *The impact of the Cognitive Skills Training on post-release recidivism on Canadian federal offenders*. Research Report R-41. Ottawa, ON: Correctional Service of Canada.

¹⁷ Serin, R. (2001). Improving program performance among offenders. *Forum on Corrections Research*, 13(1), 24–26.

¹⁸ Op. Cit., Los el (2001).

¹⁹ Preston, D.L. (2000). Addressing treatment resistance in corrections. In Motiuk, L.L., & Serin, R.C. (Eds.), *Compendium 2000 on Effective Corrections*, Vol. I (pp. 47–55). Ottawa, ON: Correctional Service of Canada.

²⁰ Wild, T.C., Newton-Taylor, B., & Alletto, R. (1998). Perceived coercion among clients entering substance abuse treatment: Structural and psychological determinants. *Addictive Behaviors*, 23, 81–95.