Guidelines for asking the right questions and communicating results

by Gerry Gaes[1]
Office of Research, Federal Bureau of Prisons, United States of America

Why, what, where, who and how are the key questions that must be asked to conduct a program evaluation. **Why**, *is the most fundamental question regarding an evaluation. The question addresses the reason an evaluation is conducted and the intended goals of the program being evaluated. In asking* **What**, *one must define the precise nature of intervention, the social and/or psychological mechanism that are to be affected, the nature of outcomes, and the program setting. The* **Where** *of program evaluation concerns the location of the program and the timing in relation to the chronology of the offenders' correctional career.* **Who**, *refers to the program participants and their characteristics. This question is important in deciding what level of generalization is made after the evaluation is conducted. The* **How** *of program evaluation refers to both quantitative and qualitative methods of evaluation. This article presents these fundamental questions and also touches on the issue of the effective communication of results.*

The Why of program evaluation

Even though this is the most fundamental question regarding an evaluation, it is probably the least likely to be addressed and the least understood. When an administrator asks for an evaluation, it is very important to get an understanding of what he or she wishes to accomplish. Too often these questions of purpose or goals are not asked. An evaluation is conducted. The results are presented and the administrator protests, "That is not what I wanted to know."

Policymakers, administrators, program designers, often do not know how to articulate their interests in what an evaluation will achieve. Thus, the evaluator must make sure that he or she understands what is being asked. This may seem a trivial point when it comes to program evaluations for correctional interventions. Surely we know that the aim of the program is to address some deficiency of the offender and to assist his or her re-integration. But these goals are often too vague. One policymaker may have in mind that for a program to be successful a large proportion of program participants must show dramatic success. An administrator may have in mind that the program will probably only help some offenders and not others and our expectation should not be too high. Some administrators are interested in knowing how to improve a program. A program designer may think that a program success is

achieved if the participant changes his or her attitudes about wanting to change their behaviour; yet this may be too low an expectation from the administrator's point of view.

Thus, the evaluator must be able to define the goals of the study, articulate measures or criteria that will satisfy the interested parties and get the stakeholders concurrence either that the research will address their concerns or that some questions will have to await further inquiry. This is best done prior to the research design and before the implementation of the program, particularly for those programs that are new or innovative. If a program is ongoing, it is still important to clarify the administrator's goals.

Rossi and Freeman[2] have devoted an entire chapter on "The Social Context of Evaluation." In that chapter they discuss the implication of evaluation, stakeholders, and the political process involved.

To give these concepts substance, I use one of the most highly charged and politically sensitive areas of inquiry that currently exists in corrections — the effectiveness of prison privatization. In one sense, the total operation of a prison can be viewed as the broadest of program interventions. In fact, there are those that argue that the ultimate judgement of prison privatization depends on whether the industry is capable of doing a better job of reintegrating the offender back into society.

Privatization is a case where the competitors are easy to identify and where the consequences of any evaluation will be hotly contested. The stakeholders consist of policymakers and decision makers (legislators and high ranking government officials). The program sponsors are either private corrections companies or government officials advocating privatization. Evaluation sponsors are typically consulting firms or universities with foundations that do outside consulting. The target participants are the inmates assigned to a particular prison or program. The program management team is composed of corporate CEO's and administrators. The program staff are all those who are hired to deliver services. The program competitors are those companies who competitively bid to deliver a program and in some cases the competitors may be public sector employees. The contextual

stakeholders are not only the individual private companies but also public labour unions and public prison administrators or legislators who line up on both sides of the issue.

Once the goals and purposes of a program evaluation have been defined, the stakeholders identified and the political context recognized, the next step is to analyze all of the components of the program and the nature of the change mechanisms that the program is supposed to address.

It is crucial to understand that the evaluation has a political context and that the results of even a well-conducted evaluation may have little or no impact on policy decisions given the political power of the various stakeholders. The proper role of the evaluator is to conduct a well designed study; to address as many of the questions that stakeholders are interested in; and to report findings and the limitations of the conclusions. Rossi and Freeman cite Donald T. Campbell's proposal that evaluators should act as the servants of "the Experimenting Society".[3] Campbell thought that the proper role of the evaluator is to report one's findings rather than to advocate for a particular program of policy. He also cautioned against a lack of humility in presenting findings.

The What of program evaluation

There are many considerations at this stage. One must define the precise nature of the intervention, the social and or psychological mechanisms that are to be affected, the nature of the outcomes, and the program setting. Rossi and Freeman advocate the development of an impact model. This is "an attempt to translate conceptual ideas regarding the regulation, modification, and control of behaviour or conditions into hypotheses on which action can be based".[4] They also discuss causal, intervention, and action hypotheses. The impact model contains a causal hypothesis that outlines the nature of the problem being addressed. How does one become an alcoholic? What is the nature of drug addiction? What are the mechanisms of sexual dysfunction? The intervention hypothesis states how the intervention will affect the mechanism of dysfunction. The action hypothesis states whether the intervention is somehow different from the mechanism that caused a problem to occur in the first place. For example, if one is designing a program to teach employment skills, the causal

> The proper role of the evaluator is to conduct a well designed study; to address as many of the questions that stakeholders are interested in; and to report findings and the limitations of the conclusions.

hypothesis states that certain skill sets and competencies are necessary to become employed. The intervention hypothesis says that vocational training will improve the set of skills; however, the action hypothesis says that while vocational training improves skills, it does not address all of the competencies required for successful employment. Other competencies include the ability to get along with co-workers or the ability to listen and take orders.

The Where of program evaluation

The where of program evaluation concerns the location of the program and the timing in relation to the chronology of the offender's correctional career. Program location may seem unimportant; however, it can often be the deciding factor whether a program is successful or not. A residential drug abuse program located in an environment where drugs are readily accessible or where staff, other than the program staff, are not supportive of the intervention is unlikely to succeed regardless of how well the program is designed. Program support is something that is not typically documented by program evaluators. This can have grave consequences for program success.

The Who of program evaluation

Defining program participants is as important as defining the nature of the program. In some cases, the characteristics of the program participants may be so important that the evaluator will want to experimentally manipulate the relation between the intervention and the target population. The risk principle is a global statement of the nature between interventions and the program participants. It says that regardless of the nature of the program or the intervention, the program will demonstrate a greater success for those offenders who are at higher risk. There are of course many other characteristics of the target population that could affect the inferences to be made. Are there gender-specific types of interventions? Are there socio-economic factors? What types of interventions have the target population participated in before? All of these questions are necessary not only to control for background characteristics of the population. They are important in deciding what level of generalization we want to make after the evaluation is conducted.

The How of program evaluation

Quantitative versus qualitative approaches

Most of the modern research on program evaluation emphasizes quantitative methods to determine whether an intervention has been successful. I am an advocate of quantitative research because I think it is the only way that the social sciences will be able to establish laws about human behaviour. But there is a great deal of room for the qualitative analyst in the social sciences and in evaluation research. Even though we assume that interventions are based upon the best science available and we may be simply expanding on an intervention that has been used before, a great deal can be learned by participant observation, interviewing program participants, or simply observing program participation with an open mind set. Anyone who has conducted serious quantitative analysis knows how much variability there is to the human response. Some of this variability may be explained by a host of variables that we use to analyze the data. But there will almost always be a great deal of residual variance. One way to approach that quantitative phenomenon is to use qualitative methods to explore the differences in human responses. Using this approach, qualitative methods are complementary to quantitative techniques.

Complementing quantitative with qualitative information

I borrow several examples from Patton's book on qualitative methods[5] to show how qualitative evaluation can be used to supplement quantitative analysis. Patton describes an evaluation of a literacy program where the evaluators used quantitative methods to measure the gain score in literacy and scales to assess participants' satisfaction with the program. While students did show positive gains from the program, the evaluators dug deeper and used individual case examples to explain the nature of the gains and open-ended interviews to enhance their understanding of satisfaction with the program.

When program participants were asked to describe their opinions about the program, they gave specific reasons why they were so satisfied. No longer constrained to the specific responses in the satisfaction questionnaire, participants described how they could now read the newspaper; make a shopping list; understand the instructions on their medicine bottles; navigate city streets better; and, how they could take the written test for their drivers license.

Qualitative data is not simply an exposition of quantitative data, it often suggests that the categories we choose to uniformly measure a phenomenon may not be the "phenomenology" of the participant. Open-ended interviews or open-ended items allows the participant to express attitudes, opinions, or beliefs that may provide a fresh insight into the program impact. This may be especially important during the early phases of a program design or implementation.

Appropriate use of qualitative methods

Patton has also outlined "Particularly Appropriate Uses of Qualitative Methods".[6] The following briefly describe each of these.

Process studies and process evaluations

Process evaluations examine the nature of how an outcome is achieved. Program evaluations should always be based on theory that articulates how an intervention will modify human behaviour. To understand the mechanism of change, the researcher can supplement quantitative measures of mediating outcomes with interviews that probe the client on the nature and causes of his or her behaviour. It is my experience that even in successful intervention programs, attempts to quantitatively relate process to outcome typically have limited success. In quasi-experimental designs or observational studies, it is particularly important to rule out artifactual or unintended causes of an outcome. Process evaluations not only examine the mechanisms of changes but the change agents themselves. Thus, program providers are also under study in a qualitative process evaluation. Patton lists the following questions: "What are the things people experience that make this program what it is? What are the strengths and weaknesses of the program? How are clients brought into the program and how do they move through the program once they are participants? What is the nature of staff-client interactions?"[7]

Formative evaluations for program improvement

Formative evaluations are intended to improve a program. These are also process evaluations that emphasize the strengths and weaknesses of a program. A program may be well-designed, based on sound theory, and well measured; yet, there may be internal group or individual dynamics that interfere with program progress. Perhaps staff are not well trained or they are not "connecting" with the clients. Formative process evaluations seek to uncover these problems.

Evaluating individualized outcomes

The matching of treatments and program services to the needs of clients is the mantra of many social workers, psychologists, and educators. Yet, matching is rarely an explicit part of a program assessment process. One way to approach matching is to do qualitative studies in which the researcher provides descriptions of the different ways clients react to different treatments, treatment styles, and treatment providers. Evaluators document the unique perspectives of clients to the treatment regimen. This may lead to a typology and eventually to a quantitative assessment of specific matching hypotheses.

Case studies to learn about special interest, information-rich cases

Cases can be chosen that represent particularly incisive information about a particular program. Perhaps case studies of extreme program failure are relevant. Structured interviews with these clients may indicate alternative strategies for subclasses of individuals. Such inquiries may extend to dropouts, or to people who show dramatic gains from a program. In each case, the researcher is interested in understanding the nature of failure or success so that the program can be improved.

Comparing programs to document diversity

When one tries to adapt a national program or a "universal intervention" to a specific location, there are many reasons to expect that there are local nuances in program implementation or potential differences in the clients. These differences may contribute to unexpected outcomes. These differences can be documented both quantitatively and qualitatively.

Implementation evaluations

The best interventions will fail if attention is not given to the implementation of a program. Most evaluators using objective, quantitative data go about their measurement of outcomes assuming that the program has been successfully implemented. There are quantitative methods to assess program implementation; however, qualitative methods can also be of assistance here. Patton addresses the problem with the following qualitative dimensions: "What do clients in the program experience? What services are provided to clients? What does staff do? What is it like to be in the program? How is the program organized?"[8] This qualitative approach should be supplemented with tests of what the client has learned or ratings of the effectiveness of the treatment provider by other knowledgeable people. Thus, once again we can complement one type of information with the other.

Identifying a program's or organizations's theory of action

According to Patton, a theory of action relates program inputs and actions to outcomes. This sounds very much like a well articulated theory. However, citing Argyris[9] Patton discusses "espoused theories" from "theories-in-use. The former are those principles advocated by program designers or program theorists. The latter are the beliefs of the treatment provider, the street level bureaucrat actually doing the work. A qualitative assessment of both will indicate the extent to which there is parallelism in the plans of the treatment designer and the treatment provider. This may be especially crucial in a new groundbreaking approach.

Focusing on program quality or quality of life

Patton argues that even if a program evaluation can be clearly defined and measured in a quantitative way, it is still important in many cases, to assess the texture and contours of meaning of program impact by doing a qualitative assessment as well. For example, if we find that an offender is less likely to use drugs after a drug treatment program, what else does this imply about the offender's quality of life? A qualitative response may add insight into the nuances of different responses given by people. What does it mean to be somewhat satisfied as opposed to be completely satisfied with one's treatment?

Documenting development over time

Developmental changes are extremely important in analyzing human and organizational growth (decline) over time. While quantitative data may indicate developmental changes are occurring, qualitative inquiry may give greater insight into the growth process. When we measure growth, we often use linear or sometimes non-linear patterns to demonstrate growth has occurred. But these may be idealized growth curves. Growth may represent sudden transitions in states for some individuals or organizations and slow or little growth in others. Trying to ascertain the growth phenomenon through qualitative analysis may provide a greater understanding of the processes under consideration.

The How of quantitative evaluation and communicating results

The how of quantitative evaluation could cover volumes. It involves research design, quantitative methods, measurement theory, meta-analysis, decisions about cost-benefit procedures, simulations, and many other technical areas. It involves precise operational definitions of the program intervention, the processes it is intended to change, and the outcomes of interest. The skill sets of the evaluators should also to be considered. Psychologists,

sociologists, economists, operations researchers, and computer simulation experts all can bring different perspectives to the evaluation approach. The few comments I want to make here relate to communicating the results of the quantitative analysis.

In their concluding chapter, "The Social Context of Evaluation"[10] Rossi and Freeman discuss the need for evaluators to become "secondary disseminators." Most evaluators are quite good at producing a technical report on the results of the evaluation. These reports are usually only read by peers and not by the stakeholders who are most affected by the

evaluation results. Thus, secondary dissemination refers to the communication of research results to the stakeholders in ways that they can understand and that are useful to making further policy decisions. This kind of communication should be direct and short. It should provide any necessary qualifications or limitations of the study, often missing from executive summaries. It should also use language that the stakeholders can understand omitting the technical jargon of the discipline. It can be a humbling experience to ask your audience what they learned from your presentation. But it is also my experience that getting their feedback is better than their silence. ■

1 320 First Street NW, Washington, DC 20534.

2 Rossi, P.H. and Freeman, H.E. (1993) *Evaluation: A Systematic Approach, 5th Edition*, Newbury Park, CA: Sage.

3 Campbell, D. T. (1991) "Methods for the Experimenting Society," Evaluation Practice. 12, no 3.

4 Rossi and Freeman, p. 119.

5 Rossi and Freeman, p. 119

6 Patton, M.Q. (1990) *Qualitative Evaluation and Research Methods: Second Edition,* Newbury Park, CA, Sage.

7 Patton, Chapter 4, p. 92-141.

8 Patton, p. 95.

9 Patton, p. 105.

10 Argyris, C. (1982) Reasoning, Learning, and Action: Individual and Organizational. San Francisco: Jossey-Bass.

New Release

The Safe Return of Offenders to the Community

Statistical Overview
April 2000
Research Branch, CSC

TABLE OF CONTENTS

To access this document, visit the CSC Web site at http://www.csc-scc.gc.ca/text/faits/facts08_e.shtml