MANAGEMENT-ORIENTED CORRECTIONS EVALUATION GUIDELINES

U.S. DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE

MANAGEMENT-ORIENTED CORRECTIONS EVALUATION GUIDELINES

By

Jack Reynolds, Ph.D.

APRIL 1979

U.S. DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE

Law Enforcement Assistance Administration
Henry S. Dogin, Acting Administrator

National Institute of Law Enforcement
and Criminal Justice
Blair Ewing, Acting Director

This project was supported by a grant awarded to the Center for Human Resources, University Research Corporation by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

TABLE OF CONTENTS

TABLE OF FIGURES

ABSTRACT

Program evaluation is readily accepted as a part of good management. But managers have found that good evaluation is not easy to come by.

Management-Oriented Corrections Evaluation Guidelines is a "how-to" manual that deals specifically with a management approach to evaluation. It was written as a reference manual for corrections administrators and evaluators who participated in a series of two-and-one-half day workshops in "Management-Oriented Corrections Evaluation."

The success of these workshops and the usefulness of the guidelines on which they were based has demonstrated that evaluation can be an effective instrument for managing and improving the correctional subsystem of criminal justice.

The guidelines describe a generic evaluation process that can be easily adapted to a wide variety of corrections programs. They are a useful resource for corrections administrators and evaluators alike.

PREFACE

In recent years the pressure for evaluation has increased at all levels of the corrections field. In response to this pressure, the Law Enforcement Assistance Administration in 1973 set up an Evaluation Policy Task Force that examined the evaluation question and recommended a program designed to:

1) Increase knowledge about the effectiveness of criminal justice programs and practices
2) Build evaluative information into program management
3) Develop evaluation capability in criminal justice agencies at the state, county, and municipal levels.

LEAA accepted the recommendations and established an evaluation program in its National Institute of Law Enforcement and Criminal Justice (NILECJ).

One of the many projects sponsored by this program was Stuart Adams' review of the state of the art of evaluation in the corrections field.* Adams found that correctional administrators often find research and evaluation not useful. Some reasons for this are:

- The field of correctional research is relatively undeveloped. Researchers in the correctional field have yet to design evaluation methods and systems that are appropriate and responsive to the practical operational questions, information needs, and decisionmaking requirements of correctional administrators.
- There is little systematic communication between correctional administrators, on the one hand, and correctional researchers and evaluators, on the other, about the role, function, and utility of research and evaluation in the correctional setting.
- Correctional administrators often do not fully understand evaluation or appreciate how it can be and ought to be an integral part of effective management and decisionmaking.

NILECJ decided to do something about this problem. In 1975, the National Institute awarded a grant to the Center for Human Services (CHS) to:

*Adams, Stuart, *Evaluative Research in Corrections: A Practical Guide.* Washington, D.C.: U.S. Government Printing Office, 1975.

- Develop guidelines for correctional administrators and correctional evaluators on how to make evaluation firmly and directly responsive to the normal and pressing needs of the correctional administrator.
- Bring both correctional administrators and their in-house evaluators together in two-and-one-half-day workshops to present them with these guidelines.
- Follow up the workshops with a limited amount of additional training for administrators and evaluators who decided that the guidelines would be useful and applicable to their back-home situations.

The Center for Human Services and its affiliate, University Research Corporation, have trained hundreds of administrators, program officers, and evaluators in how to design and conduct management-oriented evaluation, principally under contracts to the U.S. Department of Housing and Urban Development and the U.S. Department of Health, Education, and Welfare. The evaluation process has been applied to a variety of fields at local, state, regional, national, and international levels. NILECJ expected CHS to adapt its management-oriented evaluation process to the corrections field and to deliver 10 Regional workshops for up to 500 correctional administrators and in-house evaluators. CHS undertook the project under the direction of Dr. Geoffrey Wood. A panel of correctional research experts—Stuart Adams, Donald Gottfredson, Paul Kusuda, Nathan Mandel, Arthur Pearl, and Laurel Rans—was recruited to give guidance to CHS in its adaptation of the evaluation process and design of the workshops.

The author was responsible for adapting and writing the evaluation guidelines. A core training team composed of Dr. Wood, Ms. Margaret Neuse, Dr. Alvin Cohn, and Ms. Laurel Rans developed the workshops and conducted the training.

By the end of the grant period in 1977, over 750 correctional administrators and evaluators had been trained in the evaluation process. The positive response to the training and the continued interest in the refinements that were made in the guidelines led to this revision.

This document describes a generic approach to evaluation that can be used by administrators and evaluators alike. It is written principally for corrections personnel, but it can be, and has been, used by other criminal justice personnel as well. The document is intended to be a reference work that describes the steps in designing and conducting an evaluation in a "how-to-do-it" format. But besides procedures,

worksheets, and descriptions of steps to take, there are also discussions of key concepts, examples, and illustrations.

Over the years, many people have contributed to the development and refinement of the concepts and procedures that make up these guidelines. We are particularly grateful to our colleagues at the National Institute of Law Enforcement and Criminal Justice—Fred Becker, Bob Burkhart, Paul Lineberry, Lois Franco, and Linda Grossman. Stuart Adams, Laurel Rans, and Paul Kusuda, three of our advisors, were also very helpful. At CHS, special thanks are due to Margaret Neuse for her many suggestions for improving the guidelines, to Paul Mathless for editing the text, and to Vennette Fuerth and Marta Kelsey, who have provided invaluable support throughout the project. We also want to acknowledge the contributions of the many workshop participants whose comments, suggestions, and criticisms have been invaluable.

Jack Reynolds, Ph.D.
Center for Human Services

MANAGEMENT-ORIENTED CORRECTIONS EVALUATION GUIDELINES

I. INTRODUCTION

What Is Evaluation?

Academicians and evaluators enjoy arguing over *the* definition of evaluation. Like many other basic concepts, evaluation is one that everyone seems to understand until they begin to define it.

Most definitions seem to look at evaluation from one or more of three perspectives:

• **As a process**	How evaluation is done; the steps and procedures involved in designing and conducting an evaluation
• **As a product**	The results, findings, or judgments that are made as a result of doing an evaluation
• **In terms of its purpose**	The end use of evaluation—for example, for planning, policy-making, or decision-making.

Our view is that evaluation is a process that results in a product that has a purpose. That *process* is what will be described in this manual. It consists of a series of steps for determining what will be evaluated, and for planning, designing, and conducting an evaluation.

The evaluation process is essentially the same as the process that is used in research. Evaluation and research use the same techniques—sampling, data collection, analysis techniques, experimental and non-experimental designs, and so forth. What distinguishes research from evaluation is the *product*. While research is conducted to gain knowledge, evaluation is conducted to make judgments. In evaluating a corrections project, not only do you want to know what it is doing, but you also want to make judgments about how well it is doing it—to what extent it is adequate, effective, or efficient.

The products of the evaluation process are judgments—value judgments. In fact, evaluation is often described as the act of placing a value on something or someone. In this broad sense, we are all evaluators. We make judgments every day about the value of different things—books, movies, politicians, actors, transportation systems, even corrections programs.

When evaluation is done informally, these judgments are often called "subjective." The type of evaluation treated in this manual is more formal. We attempt to employ rational, objective, and precise measures to make fair or "objective" value judgments about programs.

While judgment may be the principal product, the utility of

evaluation can go beyond the exercise of judgment. That is where the *purpose* of evaluation comes in. Why is a subject being evaluated? How will the results be used? *We believe that evaluation can be most useful when it is designed to provide information needed for decisionmaking.*

In the field of corrections, evaluation can be used to provide information for making political as well as administrative decisions about funding projects, modifying services, reorganizing staff, and so forth.

One way to be certain of getting the needed information is to identify the decision options and then design the evaluation to help decisionmakers choose the best option. For example, evaluation results could help a warden decide whether to continue a special inmate education program, modify it, or cancel it. When it is related to these decision options, evaluation becomes more than a mere description of a program or activity, more than a judgment about its worth. Linked to decision options, the information and judgments evaluation produces are more likely to lead to program improvement.

Therefore, as a management information tool, we define evaluation as:

> A process for making judgments about selected people, objects, and events by comparing them with specified value standards for the purpose of deciding among alternative courses of action.

The Guidelines

To make evaluation work in the way we have described it, administrators and evaluators must work together to:

- Identify priorities and select useful evaluation topics that are feasible as well as responsive to decisionmaking needs
- Design practical evaluation methodologies and procedures
- Conduct and manage evaluations effectively and efficiently.

This manual is written as a guide to help administrators and evaluators meet these requirements. In its simplest form, the evaluation process described here can be seen as a three-phase process operating within the political and managerial context of corrections decisionmaking.

In Phase I, the administrator and evaluator work together to determine exactly what needs to be evaluated, for whom, by what date, and for what purpose. Phase I and the phases that follow it consist of a series of steps that may be completed in several different ways. But when Phase I is completed, a product should emerge—a clear definition of the evaluation topic.

Phase II is the development of the technical design and the management plan for the evaluation. In this phase, the administrator and evaluator follow steps to identify the data that will be needed to make judgments about whatever is being evaluated. They also determine how to collect and process data, and how to schedule, staff, and budget the evaluation.

Phase III is the implementation phase. The steps in this phase deal principally with carrying out the technical and management plan and producing the results.

Once the evaluation is concluded, the findings and judgments are communicated to the appropriate decisionmakers, who combine them with other information and judgments from outside this formal evaluation process. Eventually, a decision is made, an option is selected, and action is taken. That action may be to reorganize a project or distribute funds, or, in some cases, it may even be to plan another evaluation study. And the process begins again.

Figure 1 illustrates this flow and the three phases of the evaluation process.

Figure 1

THE EVALUATION AND DECISION-MAKING PROCESSES

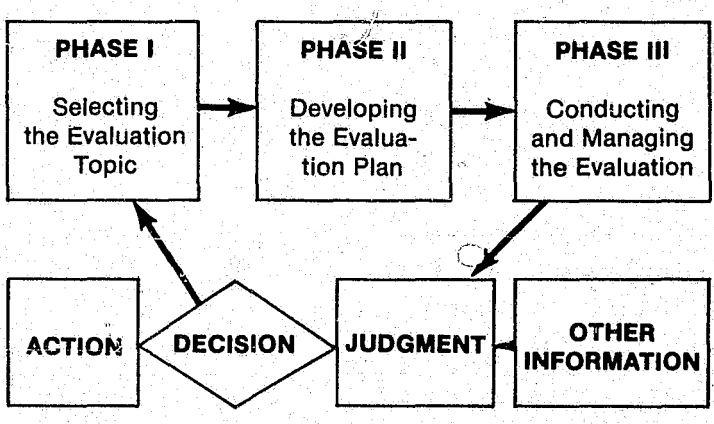


Figure 2 outlines the steps in each of the three phases. Each of these steps is discussed in detail in the following chapters.

Figure 2
PHASES AND STEPS IN THE EVALUATION PROCESS

PHASE I: SELECTING THE EVALUATION TOPIC

1. Identify the evaluation priorities
2. Clarify the evaluation subject
3. Clarify the objectives of the subject
4. Specify the objective of the evaluation
5. Ascertain the feasibility of the evaluation

PHASE II: DEVELOPING THE EVALUATION PLAN

1. Select the study design
2. Select the evaluation criteria
3. Select the evaluation standards
4. Develop the analysis plan
5. Develop the sampling plan
6. Develop the data collection plan
7. Develop the reporting plan
8. Develop the managment plan

PHASE III: CONDUCTING AND MANAGING THE EVALUATION

1. Make staff assignments for the evaluation
2. Develop the evaluation and management procedures
3. Pre-test and revise the evaluation procedures
4. Collect and analyze the data, and report the results
5. Develop strategies for using the evaluation findings

II. PHASE I: SELECTING THE EVALUATION TOPIC

Evaluation is seen as a threatening prospect by some and a golden opportunity by others. Different people want to evaluate different

things for different reasons. If an evaluation is going to be useful to these different parties, then there must be some meeting of the minds before the work begins. As David Twain, et al., stressed:

> There must be some "pay-off" for each of the major participants if collaboration is to be successful, and these immediate and potential benefits must be recognized at the outset.*

The steps in Phase I are designed to help the major participants in an evaluation—particularly the administrator and the evaluator—select evaluation topics that will have a pay-off. Because evaluation interests may vary among the participants, you may need to list these varied interests and then set some priorities. Once this is done, a specific evaluation topic can be clarified as to *what is to be evaluated, why,* for *whom,* and by *when.* At the end of this phase, the evaluation topic should be phrased in terms of a subject and an objective:

> To provide (whom) with evaluative information on (subject and type evaluation) by (date) in order to (purpose and decision options, if applicable).

Phase I consists of five steps: 1) identifying the evaluation priorities; 2) clarifying the evaluation subject; 3) clarifying the objectives of the subject; 4) specifying the objective of the evaluation; and 5) ascertaining the feasibility of the evaluation.

STEP 1: IDENTIFY THE EVALUATION PRIORITIES

1.1 Identify your evaluation interests—Consider and list questions to be answered, hypotheses to be tested, judgments and decisions to be made, program goals and objectives, problem areas, major issues

1.2 Rank order these interests according to their utility—consider utility for management, public relations, accountability, reputation

1.3 Negotiate consensus on the order of ranking (priorities)

*Twain, David; Harlow, Eleanor; and Merwin, Donald: *Research and Human Services: A Guide to Collaboration for Program Development.* New York: Research and Development Center, Jewish Board of Guardians, September 1970, p. 23.

Each of the major participants can begin by identifying the broad areas of interest he/she thinks should be evaluated. You may have your own list and separate lists for other interested parties (the public, the legislature, the evaluator), or you may combine them into one list.

In his chapter on "Research Needs and Priorities," Stuart Adams identified several sources of evaluation needs.

> One of these is the mission of the agency. Another is agency requirements for self-maintenance and self-improvement. Still another is the perceived role of corrections in relation to other agencies or systems in the community. Finally, there are such matters as correctional experience and correctional theory. These areas give the administrator his main clues as to what research is needed and what is likely to yield practical results.*

Obviously, one of the most important sources of evaluation interests should be issues related to the agency's mission: Are agency and program goals and objectives being met? Is recidivism being reduced? Is security being maintained? Are inmates being rehabilitated? Other common sources are problem areas: Why is staff turnover so high? What are the reasons for disputes among inmates?

These interest areas can be phrased in different ways. Administrators often phrase their interests as *questions* they would like answered: Is security sufficient? Is recidivism being reduced? Evaluators may think in terms of *hypotheses* to be tested: "Work release projects in rural areas are more successful than those in urban areas." Both administrators and evaluators may express evaluation interests in terms of *judgments* to be made about effectiveness and efficiency: Which is the most cost-effective service—counseling, vocational training, or recreation? You may also find evaluation interests expressed as *decisions* to be made: Should the probation workload be reduced, maintained, or increased?

Your interest areas can be identified in any form and phrased in any way. You may prefer to list them by yourself, ask for staff suggestions, poll your clients, search the literature for ideas, or think them out with your colleagues. Group involvement in this process may be a good idea. It can help ensure that the key participants in the evaluation are a part of the topic selection process, which can lead to smoother relations when the evaluation is designed and conducted.

*Adams, *op. cit.*, p. 23.

Since you will not be able to evaluate everything at once, you will have to establish some priorities. Adams wrote:

> In principle, research monies should be directed primarily into projects that the administrator deems important . . . Since all such research needs cannot be met simultaneously, they must be given a ranking—most important first. However, areas of greatest need are not necessarily areas of payoff . . . Consequently, evaluative studies should reflect both perceived needs and perceived possibilities of achieving practically important results.*

One way to establish priorities is to rank your list in terms of greatest utility. Think of how the results would be used. Can they be used to make decisions? If so, that is a higher order of utility than using the results to improve judgments. Improving judgments is more useful than merely increasing knowledge. Consider positive uses and possible misuses in determining your ranking. Be able to explain why an area would be useful to evaluate or why it would not be. Here are some examples of ways evaluation can be used:

• **Management improvement**	Improve planning, operations, management decisionmaking
• **Public relations**	Improve community awareness, increase political support, raise central issues for public consideration
• **Accountability**	Provide required information to superiors, the legislature, LEAA
• **Reputation**	Enhance the reputation of the program, the administrator, the evaluator.

If you are the only person involved in establishing evaluation priorities, then you will have already set your priorities by ranking your areas of interest. If more than one person or group is involved, then you will have to negotiate consensus on the priorities. There are many ways to do so. They all involve communication among the parties, honesty, candor, and a willingness to see the other person's point of view. One procedure is described below:

• **Compare the ranked lists**	Each participant shares his/her list with the others.

*Adams, *loc. cit.*

• **Identify common areas of interest**	Identify those items that are at the top and at the bottom of everyone's list.
• **Negotiate consensus on the top items**	Take the few items at the top of everyone's list and discuss the utilities and drawbacks of each, with all participants explaining how and why an item would be useful or not useful to them or their group. Ask the participants to reorder the list, if necessary. Close with agreement.
• **Negotiate consensus on the bottom items**	Use the same process as for the top-priority items.
• **Attempt to negotiate consensus on the remaining items**	Use the same process.

Not every item on every list has to be negotiated. Since there will probably be more items than you have the time and money to evaluate, one or two top-priority items will usually be enough to begin with.

Throughout the evaluation process, there may be several points where you will need to negotiate. The procedure just described can be used to negotiate consensus each time it is needed.

STEP 2: CLARIFY THE EVALUATION SUBJECT

2.1 Review the relevant literature, documents, records
2.2 Identify the object, person, or event to be evaluated
2.3 Determine the number of subjects to be evaluated
2.4 Identify any features to be given special emphasis
2.5 Identify the type of evaluation to be conducted: need, design, performance, impact.

Once an evaluation interest has been selected, many evaluators find it helpful to collect and review the relevant literature. For example, if the administrator is interested in knowing whether the work-release program is meeting its objectives, the evaluator might

collect annual reports, minutes of meetings, grant applications, and other documents from the files. In addition, he/she might review journal articles and books on work-release. This review can help give a context to the evaluation that will make it easier to clarify the evaluation subject.

Next, you must clarify the *objects, persons, or events* to be evaluated: a halfway house, a work-release program, the laundry service. An evaluation can be of:

- **Objects** — Cells, food, textbooks, clothing, dining halls—that is supplies, equipment, or facilities
- **People** — Inmates, guards, wardens, evaluators, volunteers, staff, clients, or professional associates
- **Events** — Programs, projects, activities, procedures, or even tasks—for example, a juvenile diversion program, a skills training project, a work-release transportation service, an inmate counseling procedure, a filing task.

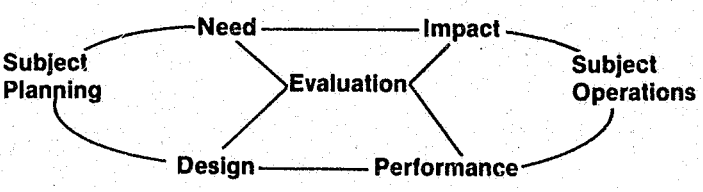
You also need to clarify the *number* of subjects to be evaluated: one recreational project, several, many? Then specify any features or parts of the subject that should be given special *emphasis*. For example, if you plan to evaluate your food services (an activity), you may want to pay particular attention to some of the procedures, such as menu preparation, serving, or security procedures. Or you may want to emphasize costs, inmate satisfaction, or speed of service.

You also should specify the *type* of evaluation to be conducted. A subject can be evaluated as to: 1) the need for it; 2) its design; 3) its performance; or 4) its impact.

Each subject can be viewed as having several stages in its life cycle, as in Figure 3 (see page 12). By clarifying the stage to be evaluated, you define the *type* of evaluation to be conducted.

You may want to know if there is a *need* for a new prison; or whether a *design* for a new prison is acceptable; whether a new prison *performs* as expected; or whether a new prison has had any *impact* on escape rates. Sometimes, you may want to have several such kinds of information at the same time. In that case, you will evaluate your subject from several perspectives simultaneously—you will be conducting several types of evaluation.

Figure 3
STAGES IN A SUBJECT'S LIFE CYCLE AND CORRESPONDING TYPES OF EVALUATION



Evaluation of Need

This type of evaluation is sometimes called *needs assessment*, and it is usually conducted when the real need for a subject is unknown. If your evaluative questions are like the following, you probably want a needs assessment:

- Do we need another staff training program already?
- Are institutions really necessary?
- Which do we need the most—the gym or the dairy?

You usually evaluate need by comparing a current situation with some standard, such as the ideal situation or what is required. For example, currently 100 of 200 inmates do not have proper shoes. Ideally, all 200 inmates should have good shoes. Therefore, our assessment of the need is:

$$\frac{\text{Current: } 100}{\text{Ideal} \quad : 200} \times 100 = 50 \text{ percent}$$

Only half the ideal is being met, or, there is a need for 100 more pairs of shoes.

Needs assessments are common in educational and training programs. To determine what things the professional staff needs to learn, some programs first conduct a functional job analysis to identify what the staff members are required to do. By comparing what they have to do with what they can already do, the educator can determine what they need to learn.

In corrections, you have many situations calling for an assessment of need—the need for people (staff, volunteers, evaluators) and the need for events (counseling services, filing, vocational training).

Although needs assessments are important, they often take a good deal of time and effort to conduct, particularly when the subject is

large, or philosophical. Evaluating the need for rehabilitation services in institutions, for example, could be a major study. But needs assessments can be small and still be useful. You could assess the needs of the guards for training, the needs of the inmates for recreational and health services, the needs of the probation officers for assistance. Needs assessments of manageable size can provide useful information for decisionmaking.

Evaluation of Design

This type of evaluation is sometimes called *plan review* or *proposal review*. It is usually conducted when a plan for some subject has been developed and formally submitted. If your evaluation questions are like these, then you may want a design evaluation:

- Is this training curriculum adequate?
- We have to determine which of these 20 proposals for demonstration projects should be funded.
- Is this design for the new dining hall acceptable?
- Will this plan for post-release counseling produce anything?

You usually evaluate the design of a subject by comparing what is proposed to what is needed. For example, if the new dining hall needs to serve 150 people per seating and the architect's plan allows 200 people to be served per seating, then the evaluation would be:

$$\frac{\text{Planned: } 200/\text{seating}}{\text{Needed: } 150/\text{seating}} \times 100 = 133 \text{ percent}$$

Evaluation shows that the plan is more than adequate with respect to seating capacity.

Design evaluation is common in the commercial sector, where contracts are open for competitive bidding. Competing firms submit proposals for "meeting the need," and these are evaluated to select the most effective and cost-effective design. In government work, grants and contracts are often awarded the same way. In corrections, you will have occasion to evaluate grant and contract proposals, as well as project and department plans. For example, many departments of corrections evaluate the plans submitted to them by their operating agencies; many SPAs also evaluate annual work plans.

Plan evaluation is important because performance evaluation often depends on the existence of a workable plan. If the plan is inadequate, it may have little or no practical utility. Thus, you should seriously consider evaluating each project's plan, whether it be a vocational education project, a work-release program, or a design for a new pre-sentence investigation scheme.

Evaluation of Performance

This is the most common type of evaluation, and in many programs it is the only type conducted. If your evaluation questions are like the following, you probably should conduct a performance evaluation:

- Did the diversion program meet its objectives?
- Which service is more cost-effective—the institutional or the community-based?
- What did participants learn as a result of this training?

Performance evaluation includes measuring the conformity of the subject to its design; the production of the expected goods and services, in terms of both quantity and quality; and the immediate effects of those goods and services on the knowledge, attitudes, motives, and behavior of the recipients.

Performance evaluation can consist of all or some of the following subtypes:

- **Conformity** — Program performance is compared with legal requirements, regulations, or federal "standards." For example, does the halfway house operation conform to Equal Employment Opportunity regulations?
- **Products** — Here you determine whether the expected products are actually being produced. Does the new refrigeration system preserve foods the way it is supposed to; does the farm produce the expected quantity and quality of produce; is the counseling program serving people according to plan?
- **Effects** — If the goods and services are being produced according to plan, the next question is: Are they having the effects expected? Are inmates learning new skills in the Voc Ed course? Are the attitudes of the guards improving due to the staff development program? Does the community support the diversion program? Effects usually refer to changes in knowledge, attitudes, motives, and behavior that are a result of the subject.

You usually evaluate performance by comparing what is actually produced with some standard, such as the subject's planned objectives. For example, say 46 parole officers were trained. The course was designed to train 50. Our evaluation of performance is:

$$\frac{\text{Actually trained : } 46}{\text{Planned objective: } 50} \times 100 = 92 \text{ percent}$$

In meeting 92 percent of its objective, this parole officer's training course can be said to have been effective.

Evaluation of Impact

This is the evaluation of the subject's more permanent, or long-range, effects on the status of the individual, group, or community. Common impact concerns in criminal justice are recidivism and crime rates. If your questions are like the following, you are probably concerned with impact evaluation:

- Did the work-release program reduce crime?
- What long-term skill gains accrued from the training?
- Which program did the most to increase inmate employability—the counseling or the Voc Ed?

You usually evaluate impact by comparing the "before" status with the "after" status. For example, before the halfway house was opened, the recidivism rate was 40 per 100 after one year. After a year of halfway house operation, the rate dropped to 20 per 100. Our assessment of the impact is:

$$\frac{\text{Status after : } 20}{\text{Status before: } 40} \times 100 = 50 \text{ percent}^*$$

If we can assume that this was the only factor contributing to the change, we would say the program was very effective, that it cut the recidivism rate in half.

For practical reasons, many programs cannot conduct all types of evaluation. Most programs have a limited amount of time, money, and personnel to devote to evaluation. Therefore, two types of evaluation—needs assessment and impact evaluation—will often not be feasible, except as special studies. These types of evaluation should usually be conducted on a regional or national level by highly trained researchers, with the results fed back to the operating level for general guidance in planning.

*Note that a figure below 100 percent can indicate ineffectiveness in some cases and effectiveness in others, depending on the manner in which the equation is constructed in each case.

Thus, clarifying the subject means defining the objects, people, and/or events to be evaluated, how many, what is to be emphasized, and the type of evaluation to be conducted.

STEP 3: CLARIFY THE OBJECTIVES OF THE SUBJECT

Whatever subject and whatever type of evaluation you have identified, you will usually want to know whether the goals and objectives of that subject are being achieved. If you are conducting a needs assessment for recreation program facilities in an institution, you will want to assess the need in relation to the objectives of the institution's programs. If you are evaluating the design of an object, such as a classroom, you will want to know the purpose of the facility. If you are evaluating the performance of a physical education instructor, you will want to know what he/she is supposed to accomplish. Usually, you will be interested in evaluating the performance or impact of a program or project, and you will want to know whether the program or project is achieving its goals and objectives.

Therefore, you need to clarify what the goals and objectives are. Objectives should be stated as *results*, not as activities. For example, what is the auto mechanics training program supposed to accomplish—increase the employability of the inmates; produce skilled mechanics? Ideally, the objectives to be evaluated should specify the knowledge, skill or behavior to be achieved, the magnitude or quantity of the change expected, and the target date for achieving the objectives. For example, in a performance and impact evaluation of an auto mechanics training program, the following objectives might be selected for evaluation:

- **Performance objective** — At least 95 percent of the inmates who complete the training will be able to pass the General Motors qualifying examination for auto mechanics.
- **Impact objective** — At least 75 percent of the participants in the auto mechanics training who are paroled will find and keep an auto mechanics job for at least six months after release.

Clarifying program objectives will help you specify even more

precisely *what* is to be evaluated. If others are involved in the evaluation, you will have to make sure that everyone has the same perceptions of what the program is supposed to accomplish. If perceptions differ about the *real* objectives of a program, this will be another point where it will be important to negotiate consensus.

STEP 4: SPECIFY THE OBJECTIVE OF THE EVALUATION

4.1 Identify the target audience of the evaluation
4.2 Define the purpose of the evaluation in relation to the target audience: information, judgment, decision-making
4.3 Specify the decisions and the decision options the evaluation should address
4.4 Determine the deadline(s) for the evaluation and the frequency of conducting it.

Next you need to answer the who, why, and when questions. Who is the target audience for this evaluation—legislators, administrators, professional staff, clients, social scientists, the public? If there are multiple audiences, such as the warden and the public, you need to note that.

Answering the why question will require some care. What is the purpose of this evaluation? Why is it being conducted? Is it only to provide information? How will that information be used? What decisions and actions might be taken on the basis of the evaluation?

To specify the purpose of the evaluation, you should think of how the target audience will use the results. Does the audience just want information? Does it want the information to make a judgment or to test a hypothesis? Does it want to end up making a *decision*—selecting a course of action?

You can formulate the purpose in terms of statements to be addressed by the evaluation:

USE	*EXAMPLES*
• **Information only**	1. Determine how much time counselors spend with clients. 2. Determine the cost per offender for the project.

• **Information for making judgments**	1. Determine whether the post-release job-counseling project is effective and efficient—is it meeting its objectives? 2. Determine which plan will provide the most security at the least cost.
• **Information for judgments phrased as hypotheses**	1. Volunteers are more effective than parole officers in identifying job opportunities for ex-offenders. 2. Halfway houses for juveniles are more efficient (in terms of cost per offender) than detention centers.
• **Information for judgments that lead to decisions**	1. Determine whether volunteers or parole officers should have primary responsibility for identifying job opportunities for ex-offenders. 2. Determine whether the probation officers' caseloads should be expanded, maintained at their present level, or reduced.

If the purpose is decisionmaking, you should specify the decisions to be made and the *decision options* that are open. Decisions could be of several types—for example:

• **Support**	Decisions to give or withhold financial, moral, political, or other support
• **Modification**	Decisions to change a program's structure, schedule, or other characteristics
• **Inquiry**	Decisions to gather more information, to investigate further.

For each decision, the decisionmaker may have several options, or courses of action, open to him/her. For example, in deciding whether to modify a probation project, the options might be: 1) do not modify; 2) modify the staffing pattern; 3) redistribute the caseload; or 4) reduce the number of probation services.

An important part of specifying objectives is to set up the timetable for the evaluation. What is the deadline for the evaluation? When does it need to be completed? What is the frequency of the evaluation? How often will it be conducted—annually, monthly, only once?

In sum, specifying the objective of the evaluation means defining the

target audience, the deadline, the frequency of the evaluation, the purpose or use of the results, and, if the purpose is decisionmaking, the specific decisions and options the evaluation will address.

STEP 5: ASCERTAIN THE FEASIBILITY OF THE EVALUATION

5.1 Determine whether the evaluation can be conducted
- Identify any technical problems that might make the evaluation impossible: lack of data, no objectives, no way to measure impact

5.2 Determine the probability that the results will be used
- Identify the internal and external constraints that might inhibit the use of the evaluation results
- Identify the positive internal and external factors that might support use of the evaluation results

5.3 Revise the subject or objective, if necessary.

At this point you need a reality test. Is this evaluation feasible? Is it technically possible? Are there enough resources to design and conduct the evaluation? Will the results be used as planned?

To determine whether the evaluation is technically feasible, the evaluator should make a quick check of the Phase II steps. What type of study design might be required to provide the type of information needed? What criteria, standards, and measures might be needed? What might the data sources be? What kind of sample might be needed? Are there problems in any of these areas? Are data available? Can you get access to the data? Is it possible to measure, say, the impact of a program on recidivism?

You will also have to make a quick check of the resources that would be required to do this type of study. Or you might estimate what resources are available and use that information to determine what type of study would be feasible. The following are key resources:

• **Money**	Evaluation grants; budgets vs. estimated costs
• **Time**	Available time vs. quantity and duration needed
• **Staff**	Available staff and consultants vs. number and type needed

• **Material and equipment**	Available facilities, equipment, and supplies vs. those needed.

You will also have to determine the probability that the results will be used as planned. What internal or external constraints might inhibit the use of the results? What internal and external factors might support the use of the results?

For example, in evaluating the need for new facilities, you may identify constraints to one of your options—there may be no community or political support for constructing a new prison. Thus, you may decide that the decision option in this case does not really exist. The following are common types of constraints:

Internal		**External**	
Personnel	**Organization**	**Bureaucratic**	**Logistical**
Time	**Planning**	**Community**	**Physical**
Money	**Leadership**	**Cultural**	**Political**
Facilities	**Coordination**	**Economic**	**Psychological**
Records, data	**Communication**	**Geographic**	**Religious**
	Cooperation	**Legal**	**Social**

Your final step should be to revise the subject or objective to reflect the constraints. Then you should be able to define your evaluation problem as a clear and precise topic, noting exactly what is to be evaluated. One way to do that is to write out the objective of the evaluation in the following format:

To provide_____(whom)_____with evaluative information on__(subject and type evaluation)_____ by__(date)__in order to_____(purpose and decision options, if applicable)._____

Here are some examples:

To provide the Grants Program Monitor and the Legislature with evaluative information on the performance of the legal services project by January 1, 1978, in order to identify:

1. The number of inmates who use the legal services
2. The opinions inmates have about the legal services
3. The suggestions inmates and lawyers have for improving the services.

To provide the Director of Youth Services with evaluative information on the impact of the volunteer program by July 30, 1977, in order to judge whether:

1. The program is effectively meeting its objectives
2. The program is operating efficiently.

To provide the Warden with evaluative information on the performance of the work-release program by June 6, 1977, in order to decide whether the program should be:

1. Continued as is
2. Modified
3. Expanded
4. Dropped.

III. PHASE II: DEVELOPING THE EVALUATION PLAN

Once the evaluation topic has been clearly defined, a methodology and management plan can be developed. The guidelines in this phase will help you decide how to do the evaluation—specifically:

- How the evaluative judgments will be made
- How the data will be collected and processed
- How the evaluation will be organized and managed.

Although this is a more technical phase the administrator needs to be involved, particularly in deciding how the judgments will be made and how the evaluation will be managed.

A. THE JUDGMENT STEPS

The heart of evaluation is the judgment process. Given the objective that you stated for the evaluation, you must ask yourself what judgments need to be made.

Judgments can be of two types: effectiveness and efficiency. When we make judgments about something's being good or bad, adequate or inadequate, satisfactory or unsatisfactory, and so forth, these are all different ways of saying that the subject is effective or ineffective, efficient or inefficient. "Effectiveness means to produce the desired result, to accomplish the correct end, or to secure the relevant outcome. Efficiency . . . means to accomplish any stated objective without wasting resources."*

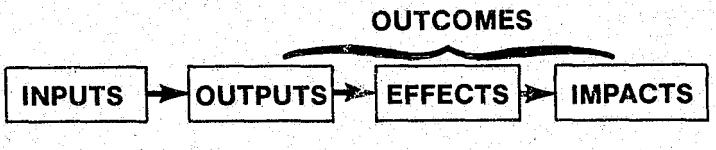
How do you construct judgments about effectiveness and efficiency? One way is to describe your program as a system of inputs and outcomes and construct effectiveness and efficiency relationships from that system. Figure 4 illustrates such a framework for a

*Hare, VanCourt, *Systems Analysis: A Diagnostic Approach*. New York: Harcourt, Brace and World, 1967, pp. 202-203.

probation program. *Inputs* are resources that are invested in a program or activity (staff, money, equipment, etc.). These are combined in various ways to produce *outputs*—the products or services of the program or activity (probation contacts or investigations). These outputs are supposed to have *effects* on knowledge, attitudes, or behavior (discharges, revocations, increased job-seeking). The effects of a program are intended to have some *impact* on the individuals, groups, or communities who are the beneficiaries of the program (increased employment rates, decreased crime). Some evaluators prefer to combine outputs, effects, and impacts and call them *outcomes*. Others do not distinguish between effects and impacts.

Figure 4

SYSTEMS RELATIONSHIPS



Probation officers	**Contacts**	**Discharges**	**Crime rates**
Pre-sentence specialists	**Caseloads**	**Revocations**	**Conviction rates**
Funds	**Visits**		
	Pre-sentence investigations		

Effectiveness can be measured by selecting an input, output, effect, or impact and making an appropriate comparison. For example, in evaluating the effectiveness of a post-release job-counseling project, one important *effect* might be the number of ex-offenders who obtain jobs. To construct a measure of effectiveness, you could compare this number with the number expected to obtain jobs. An example of how this would be done is shown below:

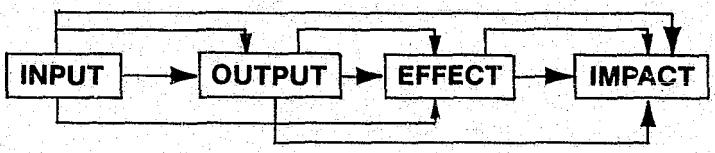
Effectiveness Criterion	Measure	Data	Percentage Effectiveness
EFFECTS			
Jobs obtained	**No. securing jobs** / **No. expected to secure jobs**	**40** / **35**	**× 100 = 114%**

In this example, five more inmates than expected got jobs; thus, the program is 14 percent more effective than expected.

Efficiency, including cost-effectiveness and cost-benefit, can be measured by selecting combinations of inputs and outputs, inputs and effects, and so forth. These criteria span the process that converts resources into products, and thus they show the relationship between the two. Several combinations are possible.

Figure 5

RELATIONSHIPS THAT SHOW EFFICIENCY



For example, an input/effect relationship might be selected to evaluate the efficiency of the post-release job-counseling project.

Efficiency Criterion	Measure	Data		Percentage Efficiency
INPUT/EFFECT (Cost-Effect)				
Cost per job obtained	No. dollars spent/ No. obtaining jobs / No. dollars budgeted/No. expected to obtain jobs	$2516/40 / $3000/35	= $62.90 / $85.71	× 100 = 73%*

In the above example, the program is cost-effective (efficient) because it uses fewer resources than expected in achieving its objective. The actual cost per job obtained was $22.81 less than planned and the program is 27 percent more efficient than planned.*

The specific measures of effectiveness and efficiency that you would select, and the number of them, would depend on the purpose of the evaluation. The four judgment steps should help you select the measures you need for the decisions you have specified.

*Again, note that a figure above 100 percent can denote effectiveness in one case, while a figure below 100 percent can denote effectiveness in other cases.

STEP 1: SELECT THE STUDY DESIGN

1.1 Review your evaluation needs (check the purpose, audience, timing, constraints)
1.2 Consider the alternatives: experimental vs. non-experimental designs
1.3 Select a design.

Before you select the specific measures of effectiveness and efficiency, you must review your needs (the Phase I summary). You also must determine the degree of accuracy and precision that will be needed for judgments. For the purpose and target audience you have identified, how accurate and precise must the judgments be? How great must your certainty be that the results are valid? Are general estimates of changes enough, or do you have to have "hard" data carried out to two decimal places?

Most administrators do not require as much accuracy and precision as do most evaluators, so the question of precision may be another issue to negotiate. There are many ways to improve accuracy and precision, but they all involve additional costs. You can select an experimental rather than a non-experimental design; you can select a large number of criteria instead of the bare minimum; you can use several standards instead of just one; you can draw a random sample instead of picking people who are more easily accessible. It's basically a matter of tradeoffs. Your result can be more accurate and precise if you are willing to pay the cost.

You can begin by reviewing the statements you formulated in specifying the purpose of the evaluation. Is the purpose to increase knowledge, improve judgments, or influence decisionmaking? Whatever the purpose, ask yourself what will be required to achieve that purpose for the target audience. That will help you assess your evaluation needs.

- **Description** — Will merely describing *what* happened be sufficient—for example, that the recidivism rate dropped 5 percent, or that 500 parolees are now employed?
- **Explanation** — Will the audience also need to know *how* and *why* the subject is effective or efficient—for example, that the work-release program failed because of a lack

of qualified staff, community support, and funds?

- **Proof** — Will the audience also demand *proof* of a cause-effect relationship—for example, that a controlled experiment showed that recidivism dropped 10 percent because of the work release program?

For descriptive purposes, you can often settle for an "extensive" non-experimental design, such as a sample survey or a simple input/output description. For explanation purposes, you will probably need more information, and you may choose an in-depth case study or a more detailed before-and-after study. If proof of causality is needed, then you will need an experimental approach.

Also think about the timing of the evaluation and the constraints that you identified. If time is short and resources limited, you will probably not be able to afford a sophisticated approach, such as an experiment.

True Experimental, Quasi-Experimental, and Non-Experimental Designs

There is no "best" evaluation design. You must select one that will provide you with the precision you require at a price you can afford. That may mean you will settle for a "quick and dirty" case study in one situation, a computer simulation in another, and a sample survey in another. Since evaluation employs the same techniques as research, you have a very sophisticated arsenal of weapons at your disposal. These can be classified in two major categories: *experimental* and *non-experimental*.

Experimental designs are the most precise investigative approach for "proving" a cause-effect relationship—for example, that community-based programs result in lower recidivism. Selltiz, *et al.*, described the logic behind experiments:

> The basic outline of an experiment is simple: an "experimental" group is exposed to the assumed causal (or independent) variable while a "control" group is not; the two groups are then compared in terms of the assumed effect (or dependent variable). This pattern makes possible the collection of the three major types of evidence relevant to testing hypotheses about causal relationships: (1) evidence of concomitant variation—that is, that the causal variable and the dependent

variable are associated; (2) evidence that the dependent variable did not occur before the causal variable; and (3) evidence ruling out other factors as possible determining conditions of the dependent variable.*

Experimental designs have been classified as *true experiments* and *quasi-experiments*. The difference between them is a matter of the degree of control over the three types of evidence mentioned by Selltiz, *et al.* Quasi-experiments do not control for as many sources of invalidity as do true experiments. Campbell and Stanley have discussed the advantages and disadvantages of 16 designs: three-non-experimental (they call them pre-experimental); three true experimental; and 10 quasi-experimental.** Adams identified nine of the most common designs used in corrections.*** These are summarized in the next few pages.

True Experiments

- **Pre-test/post-test, with control group**

  $O \leftarrow X \rightarrow O$ **Experimental group**

  $O \longrightarrow O$ **Control group**

This is the classical experimental design described by Selltiz. Both groups are observed (O) or tested. Then one group gets the experimental treatment (X). Then both groups are observed again (O) to detect changes in the experimental group that did not occur in the control group. The key to the design is *random* assignment of the subjects to the experimental and control groups. For example, in testing a human relations training course for probation officers, you would randomly assign one-half of the officers to the course, while the other half would not receive the training. You cannot assign volunteers or those who are interested, nor can you match your control group to your experimental group. Random assignment means there is no bias; every subject has the same probability of: a) being in the study, and b) being in the experimental group. This

*Selltiz, Claire, *et al. Research Methods in Social Relations*, revised edition. New York: Holt, Rhinehart and Winston, 1959, p. 94.

**Campbell, Donald T., and Stanley, Julian C., "Experimental and Quasi-Experimental Designs for Research and Teaching," in Gage, N.L. (ed.) *Handbook of Research on Teaching*. Chicago: Rand McNally & Co., 1963, pp. 171-246.

***Adams, *op. cit.*, pp. 43-73.

requirement of random assignments is often impractical, and this is why the true experiment is so difficult to undertake.

- **Post-test-only control group design**

X→O	Experimental group
O	Control group

This design also requires random assignment—and is often impractical for that reason. The only difference between this and the first design is the absence of a *pre-test.*

A pre-test is not essential to true experimental design. It provides you with a "baseline" so that you can see how much change has occurred. But the amount of change is tangential to the central question of whether X did or did not have an effect. This question can be answered without a pre-test, if the two groups are randomly assigned. An added advantage to this design is that you avoid the possible effects the pretest itself might have on the subjects.

Experiments are demanding designs that require careful planning and execution. Adams notes that the true experiment "may be losing ground to the quasi-experiment as the 'work horse' of rigorous correctional evaluation."* Two of these work horses are illustrated below.

Quasi-Experiments

- **Non-equivalent control group (matched)**

O→X→O	Experimental group
O⟶O	Control group

In this design, the subjects are not randomly assigned, thus making the design more practical. That is one reason why Adams suggests that it "may be the most useful quasi-experimental design for the correctional evaluator."** An example was a study of 110 early releasees carefully matched with 110 full-term releasees on such characteristics as age, type of offense, number of convictions, and so forth. The study, done in Florida, showed a 13.6 percent recidivism rate for the early releasees compared with 25.4 percent for the matched group.*** The matching of the subjects has to be done carefully

Op. cit., p. 60.
**Ibid.*
***Ibid.*, p. 61.

to avoid bias. The design is particularly good when the group has assembled naturally—that is, has not been brought together by the evaluator for his/her own purposes. Variations on this design have been recommended by Hatry, *et al.*,* and Weidman, *et al.***

- **Time Series** O→O→O→X→O→O→O

A series of pre-program observations are made; then the experimental variable is introduced. After the experiment, a second series of observations is made. An example would be an experimental project to close all the juvenile institutions in a state. Periodic observations of juvenile crime rates could be made, say quarterly, then the institutions shut down, then quarterly measures made again. This design is similar to the before-and-after non-experimental designs, except there are several observations before and several after. It is different from the non-experimental trend analysis in that the experiment is a one-time occurrence, not a continuing program.

Non-Experimental Designs

Non-experimental designs are more likely to be used in correctional evaluation than are experimental ones. Adams lists the following as the most popular: "the case study, the survey, the *trend analysis*, the *cohort analysis*, and the *before-after study.*"***

- **One-shot case study** X→O

Case studies are often intensive. Large amounts of data are gathered from a small number of cases. Anthropologists, psychiatrists, and business analysts often use in-depth case studies. Case studies have been conducted of team policing, youth reception centers, drug programs, and many other subjects. Case studies are probably the weakest, but also the

---

*Hatry, Harry; Winnie, Richard E.; and Fish, Donald M.; *Practical Program Evaluation for State and Local Government Officials.* Washington, D.C.: The Urban Institute (URI-17000), 1973.

**Weidman, Donald R.; Waller, John D.; MacNeil, Dona; Tolson, Francine, L.; and Wholey, Joseph S.; *Intensive Evaluation for Criminal Justice Planning Agencies.* Washington, D.C.: U.S. Government Printing Office, July 1975.

***Adams, *op. cit.*, p. 53.

most used of designs. When using a case-study design, one cannot be very sure that the experimental treatment (X) was the cause of the observed effect.

- **One-shot survey** $X \rightarrow O$

Surveys are at the other end of the continuum. Usually a limited amount of data is gathered from a large number of cases. Surveys can be of inmate attitudes, community perceptions, probation officer caseloads, or almost any other subject.

- **Trend analysis** $X \rightarrow O_1 \rightarrow X \rightarrow O_2 \rightarrow X \rightarrow O_3 \rightarrow X \rightarrow O_4$

This is periodic measurement of a subject, such as a semi-annual assessment of recidivism rates or weekly tests of inmate attitudes toward guards.

- **Cohort analysis**

Cohorts are groups of people with something in common, such as all men born in 1952, all college graduates in 1975. Corrections evaluation often studies the "release cohort" to record its performance periodically after release from a corrections program. The usual groupings are yearly releases, and they can be assessed as to the proportion who are employed, returned to the system, arrested, or any of a variety of criteria.

- **Before-after studies** $O \rightarrow X \rightarrow O$

The pre/post-test design is commonly used in corrections to measure gains in knowledge, skills, and so forth. Before-and-after studies can be easily adapted to almost any program or project. All they require is some planning to make sure the base-line data are collected before the program begins.

Another set of design possibilities can be found in operations research models.* These are usually non-experimental approaches that give you a framework for ordering different variables so that you can identify how they relate to one another. Operations research models are often used for decisionmaking in private industry. They are based on mathematical analysis, and many require computer processing. Among the most popular models are linear (for example, linear

*For a review of the use of these approaches in corrections, see Adams, *op. cit.*, Chaps. 12-14.

programming), network (for example, PERT, CPM), dynamic (for example, location analysis), and stochastic (such as the two-stage linear model).

Now that you have reviewed your needs and are aware of the principal study designs available to you, the next step is to select a design that will meet your needs and is feasible for you to undertake. Again, you will have to compromise. Elegant designs (experiments and operations research models) are usually more precise, but they also take longer and cost more. Simple designs are easier, quicker, and less expensive, but they also are less precise.

STEP 2: SELECT THE EVALUATION CRITERIA

2.1 Review your needs (check purpose, decision options, program objectives, emphasis, and constraints)
2.2 Consider the alternatives: effectiveness or efficiency or both
2.3 Select the criteria
2.4 Assign weights (unweighted, ranked, scores)
2.5 Select the measures (counts, rates, ratios, proportions, percentages, indexes, formulas).

Criteria are the characteristics, properties, or concepts that are used to make judgments about a subject. Researchers call them variables.*

What criteria do you need? You can clarify this point by reviewing five key steps: purpose, decision options, program objectives, subject emphasis, and constraints. The most important items to review are the purpose and the decision options. Look at the statements you want the evaluation to address. If the purpose is to decide among options, ask what judgments must be made about each option to arrive at that decision, and what criteria are needed to make those judgments. For example, if an option is to provide staff training in counseling, then some judgment may be needed about the

*See Adams, *op. cit.*, pp. 46-47, for a discussion of seven "widely used performance criteria" in corrections evaluation: arrests, type of offense, time until arrest, conviction, length of time in lock-up, costs of correctional treatment, and benefits.

counselors' effectiveness at present, and criteria may, therefore, be needed on inmate problems treated and not treated.

Another important source for identifying criteria is the statement of program objectives. Usually you will want to know if the program is effective and efficient in achieving its objectives. If an emphasis has been stated, such as "inmate satisfaction with grievance procedures," then you know that criteria will be needed for the emphasis areas—inmate attitudes and perceptions. The constraints may have to be included in the evaluation. For example, if you think there is community opposition to opening a halfway house, and the halfway house is one of your decision options, you may want to include community attitudes among your evaluative criteria.

As mentioned previously, there are two broad classes of criteria from which to choose: effectiveness criteria and efficiency criteria (which include cost-effectiveness and cost-benefit findings). You now can select the appropriate criteria for each decision option. For example, you might select all or some of the following as criteria for deciding whether to continue a work-release program or modify it.

Effectiveness Criteria

- **Impact**
  1. Recidivism
- **Effects**
  2. Placements
- **Outputs**
  3. Jobs developed
  4. Employers contacted
- **Inputs**
  5. Offenders approved
  6. Program costs

Efficiency Criteria

- **Input/output**
  7. Employer contacts per approved offender*
  8. Cost per job developed
- **Input/effect**
  9. Cost per placement
  10. Placements per approved offender*
- **Input/impact**
  11. Cost per non-recidivist
- **Output/effect**
  12. Jobs developed per placement
- **Effect/impact**
  13. Placement per recidivist.

*At times you may find it clearer to invert the efficiency relationships. For example, contacts per approved offender (an output/input relationship) is easier to deal with than approved offenders per contact (input/output).

Once the criteria have been selected, they should be weighted. Weighting is the assignment of a number to a criterion to indicate its importance relative to other criteria. Weighting can take several forms:

- **Unweighted** — All criteria are of equal weight.
- **Ranking** — The criteria are listed in order of importance.
- **Scoring** — Each criterion is given a score, such as cost = 5, skill gain = 3. Often, the total of the scores of all the criteria adds up to 100, just for convenience.

The last step is to select a measure for each criterion. Measurement is the assigning of a number to a criterion according to rules. There are several broad categories of measures from which to choose:

- **Count** — The simplest type of measure. For any given object or event, simply count the number, such as number of visits, number of dollars, number of favorable responses.
- **Rate** — Measures the frequency of occurrence of some event, such as miles per hour, or clients per day.
- **Ratio** — Two numbers related to each other in a fraction or decimal, such as number of guards to number of inmates (1:25 or 1/25). Any fraction, quotient, proportion, or percentage is a ratio.
- **Proportion** — A special type of ratio expressing a relationship between a part and the whole. The numerator represents a portion of the total; the denominator is the total. For example, five social workers out of 45 staff gives a proportion of 5/45, which, of course, reduces to 1/9.
- **Percentage** — If you multiply a proportion by 100, you have a percentage. Using the above proportion, the percentage of social work-

ers among the staff (5/45 x 100) is 11.1 percent.

- **Index** — A composite of two or more numbers designed to indicate a certain condition. An index of security might consist of number of attempted breaks + inmate days x 1000 (for example, 3 + 45,000 x 1000 = .06).
- **Formula** — An index is often derived through a formula. Other formulas can become very complex, involving weighted components, complicated algebra, etc.

In selecting the specific criteria that you will use, you may find it convenient to arrange them in a format such as the following example of performance criteria, measures, and weights for a work-release program.

EFFECTIVENESS	NO.	CRITERIA	MEASURES	WEIGHTS
Impacts	1.	Recidivism	No. approved offenders returned to incarceration	10
Effects	2.	Placements	No. approved offenders placed in jobs	10
Outputs	3.	Jobs developed	No. jobs developed for approved offenders	10
	4.	Employers contacted	No. employers contacted for jobs	5
Inputs	5.	Offenders approved	No. offenders approved for work-release	5
	6.	Program costs	No. dollars spent on work-release program	10

EFFECTIVENESS	NO.	CRITERIA	MEASURES	WEIGHTS
Inputs/ outputs	7.	Approved offenders per employer contacted	No. offenders approved / No. employers contacted	5
	8.	Cost per job developed	No. dollars spent / No. jobs developed	5
Inputs/ effects	9.	Cost per placement	No. dollars spent / No. approved offenders placed in jobs	10
	10.	Placements per approved offender	No. placed / No. offenders approved	5
Inputs/ impacts	11.	Cost per non-recidivist	No. dollars spent / No. not returned to incarceration	10
Output/ effect	12.	Jobs developed per placement	No. jobs developed / No. placed	5
Effects/ impacts	13.	Placements per recidivist	No. placed / No. returned to incarceration	10
			TOTAL	100

STEP 3: SELECT THE EVALUATION STANDARDS

3.1 Review your needs (see the type of evaluation and study design)
3.2 Consider the alternatives: need, demand, plan, past performance, similar subjects, control groups, requirements, professional standards, optimal model
3.3 Select the standards
3.4 Assign weights
3.5 Select the measures.

Standards are the levels or ranges of performance with which actual or proposed performance is compared. Sometimes they are called norms, normal values, or expected values. You must select a standard for each criterion. For example, you might judge a staff training program on the basis of two criteria: knowledge gain and cost. Even when you have determined that knowledge gain is 13 and cost is $250, you cannot make a value judgment about the effectiveness of the training until you have standards with which to compare these findings. You might compare staff knowledge gain, for example, with that of a control group, or with course objectives, or with knowledge gain resulting from a similar course held last year.

To determine what type of standard you should select, begin by reviewing your needs. Pay particular attention to the program objectives and the type of evaluation you identified in Phase I and the study design you selected in this Phase. Four types of evaluation were described. They usually call for the following types of standards:

• **Needs assessment**	Compare the current situation with the desired or required situation—for example, the need for training = the difference between current skills and required skills.
• **Design evaluation**	Compare the needed situation with the planned objective—for example, training program design = training needs compared with training plans.
• **Performance evaluation**	Compare the actual performance with planned performance (goals and objectives)—for example, training performance = planned skill gain compared with actual skill gain.
• **Impact evaluation**	Compare the terminal situation with initial situation—for example, training impact = post-training employment compared with pre-training employment.

The study design that you selected will also help you identify your needs. If you selected an experimental design, you will probably compare your subject with some control group, such as a group that did not receive training. If you selected a before-and-after design, you will compare actual performance after training with past performance (pre-training).

You have many different types of standards from which to choose. The following eight are the most common.*

• **Need**	Compare what is proposed with what is needed; or compare what is accomplished with what was needed—needed skills, needed food.
• **Demand**	The demands or expectations of some target population are used as the standard—inmate satisfaction, the governor's expectations.
• **Plan**	The objectives are used for comparison—Did the diversion program meet its objectives? Are the recreational objectives realistic?
• **Past performance**	A subject is compared with itself over time—How did the counseling project do last year? What is the recidivism trend?
• **Similar subjects**	One subject is compared with another that has similar characteristics—prison A with prison B; halfway house D with halfway house F.
• **Control groups**	In experimental designs, the performance of the group that gets the treatment is compared with a randomly selected group that does not—half of the inmates get a free period to do whatever work they wish; half do not.
• **Requirements**	A subject is compared with established regulations, laws, policies, guidelines, rules—A grant application is compared with LEAA guidelines, a parolee's situation is compared with parole regulations.
• **Professional**	Experts or professionals, such as physicians, judges, or parole boards, set "desirable" levels of performance.

*See Adams, *op. cit.*, pp. 44-45, for a detailed description of the more commonly used standards of comparison.

In some cases, you may select several different standards for the same criterion.

Once the standards are selected, the weights should be assigned and the measures selected as described previously. The same measures and weights should be used for the criteria and standards.

In selecting the standards, you may expand the format illustrated previously for criteria. For example:

EFFECTIVENESS	NO.	CRITERIA	MEASURES	WTS.	STANDARDS	MEASURES	WTS.
EFFECTIVENESS							
Effects	2.	Placements	No. approved offenders placed in jobs	10	Plan	No approved offenders expected to be placed in jobs	10
Inputs	6.	Program costs	No. dollars spent on work-release program	10	Budget	No. dollars budgeted	10
EFFICIENCY							
Inputs/ effects	9.	Cost per placement	No. dollars spent / No. approved offenders placed in jobs	10	Plan	No. dollars budgeted / No. approved offenders expected to be placed in jobs	10

STEP 4: DEVELOP THE ANALYSIS PLAN

4.1 Review your needs (review the purpose, study design, criteria, and standards)

4.2 Consider the alternatives: statistical or non-statistical

4.3 Outline the plan: classification, codification, tabulation, statistical manipulation, interpretation.

The analysis plan should show how the data will be treated to demonstrate the subject's effectiveness and/or efficiency. By reviewing the purpose of the evaluation, you can determine whether you need to demonstrate causality, provide a detailed explanation, or merely describe the subject. Again, refer to the statements you want the evaluation to address. The study design, criteria, and standards should also help you clarify your needs. Have you chosen effectiveness only? Are you making several comparisons? If you have selected

an operations research model, then your analytical procedures may already be spelled out in the protocol for that model.

Many analytical techniques are available to you. Some are very sophisticated and will require expert consultation; others are quite simple, and they may be adequate for your purposes.

Data can be analyzed statistically or nonstatistically. In the latter, you do not perform mathematical computations, but you still can classify the criteria in various ways and attempt to describe, demonstrate correlations, explain causality, or predict future outcomes. Statistical analysis has the same uses, but involves mathematical manipulations. Listed below are some of the more common nonstatistical and statistical analytical techniques used in corrections evaluation:

- **Description** — Narrative descriptions, frequency counts, frequency distributions, averages (means, modes, medians)
- **Correlation** — Narrative conclusions about "time-bound" associations (one criterion changes at about the same time as another—increased budgets and increased crime); rank order; correlations; linear regression
- **Causality** — Narrative conclusions about "time-ordered" associations (one criterion changes before the other—decline in food quality and prison riots); t-test; chi-square
- **Prediction** — "If-then" narrative conclusions about future events—if institutions are closed, then crime rates will increase; prediction indices.

Once you have considered your needs and alternatives, you should be ready to outline the analysis procedures you will use. Briefly, you should describe how the data will be:

- **Classified** — What broad groupings do you plan—age, type of offense, release date? If you use the input/output approach, you already have your major categories.
- **Coded** — Will the data be pre-coded or coded after collection? If you are working with large

amounts of data and with computers, pre-coded data can save you much time.

- **Tabulated** — What tables will you have, and how many? If you follow the input/output approach, your tables may already be constructed.
- **Manipulated** — What kind of statistical analyses do you plan—computation of effectiveness and efficiency ratios, tests of significance, factor analyses, etc.?
- **Interpreted** — Can you anticipate any interpretation at this point? Can you set up decision rules for your options? For example, "If the inmates do not show a skill gain of at least 15 points, the course will be cancelled."

The basic procedures for analyzing your data to determine effectiveness and efficiency were illustrated at the beginning of Phase II. If you have specified decision options, you can identify which criteria and standards will be used for which options and how the data will be interpreted. For example, the following Phase I evaluation topic has been expanded to relate purposes to analysis and action.

To provide the warden with evaluative information on the performance of the work-release program by June 6, 1977, in order to:

- **Purpose** — Decision—determine whether the program should be continued or modified.
- **Criteria and standards** — Nos. 2, 3, 4, 5, 6, and 9
- **Analysis** — Data will be classified into input, output, effect, and cost-effectiveness categories, uncoded, with a simple frequency count, ratios or percent of effectiveness and efficiency, and narrative interpretation.
- **Interpretation** — Decision rule—continue if effective and efficient on all three criteria; otherwise, modify and then re-evaluate.

At this point, it may be useful to set up dummy tables with hypothetical data to see if the criteria, standards, measures, and study de-

sign that you selected will provide you with information useful for decisionmaking. In fact, if you are using your program objectives as your standard, you can write these in now. The dummy data may stimulate the administrator to consider what he/she would really do if the program were to fall short of its objectives by 2 percent, 5 percent, or 15 percent.

Figure 6 is an Illustration of Performance Evaluation of a Work-Release Program, using real standards but dummy performance data.

The dummy data can now be related back to the original purposes of the evaluation:

> Determine whether the program should be continued. The decision rule is to continue the program if it is effective and efficient on all three criteria selected—placements, costs, and cost per placement. The dummy data show that it is efficient (cost per placement is less than planned) and effective in terms of placements. But it is not effective in terms of costs. The program is over-spending. Therefore, the decision would be to modify the program first, then re-evaluate it to see if it should be continued.

If the results are not useful, or if you are unwilling to select the option after seeing the data, then this is the time to revise the options or the judgment process.

All of the judgment procedures should be set up so that they are clearly linked to the purposes of the evaluation that were specified in Phase I.

B. THE DATA PROCESSING STEPS

Once you have established how evaluative judgments are to be made, you are ready to determine what data will be needed and how they will be processed.

STEP 5: DEVELOP THE SAMPLING PLAN

5.1 Review your needs (check the study design and criteria)
5.2 Consider the alternatives: nonprobability vs. probability samples
5.3 Outline the sampling plan: population, sample design, sample size.

Figure 6
ILLUSTRATION OF PERFORMANCE EVALUATION OF A WORK-RELEASE PROGRAM

TYPE EVALUATION: PERFORMANCE			STANDARD: PROGRAM PLAN		ANALYSES	
NO. CRITERIA	MEASURES	DATA	MEASURES	DATA	EFFECTIVENESS/ EFFICIENCY	INTERPRETATION
EFFECTIVENESS						
Effects						
2. Placements	No. approved offenders placed in jobs	140	No. approved offenders expected to be placed in jobs	≥125	$\frac{140}{125}$ = 112% 140-125 = +15	+ Exceeding objectives; very effective
Outputs						
3. Jobs developed	No. jobs developed for approved offenders	165	No. jobs expected to be developed for approved offenders	≥150	$\frac{165}{150}$ = 110% 165-150 = +15	+ Exceeding objectives; very effective
4. Employers contacted	No. employers contacted for job	205	No. employers expected to be contacted for jobs	≥200	$\frac{205}{200}$ = 103% 205-200 = +5	+ Meeting objectives
Inputs						
5. Offenders approved	No. offenders approved for work-release	560	No. offenders expected to be approved for work-release	435 ± 5% = 413-457	$\frac{560}{457}$ = 123% 560-457 = +103	- Ineffective; too many in program
6. Program costs	No. dollars spent on work-release program	$269,412	No. dollars budgeted	≤$249,500	$\frac{269,412}{249,500}$ = 108% 269,412 -249,000 = +19,912	- Overspending; ineffective
EFFICIENCY						
Inputs/effects						
9. Cost per placement	No. dollars spent / No. approved offenders placed in jobs	$\frac{\$269,412}{140}$ = $1924.37	No. dollars budgeted / No. approved offenders expected to be placed in jobs	$\frac{\$249,500}{125}$ = ≤ $1996.00	$\frac{1924.37}{1996.00}$ = 96% 1924.37 -1996.00 = -71.63	+ meeting objectives; efficient; cost per placement less than planned

≥ means greater than or equal to
≤ means less than or equal to
± means plus or minus

Sampling is the procedure by which you can draw conclusions about a large body of subjects (a population) from information that you gather from only a few of them (a sample). Sometimes you will do a case study. Then you will sample only one element of the total population. Sometimes you will look at the entire population, a 100 percent sample or a census. Anything in between is a sample. Your study design will help you identify what needs to be sampled and how precise you have to be. Experimental designs will require more careful sampling than non-experimental designs.

Your subjects and criteria will also suggest your sampling needs. For example, if you are assessing probation officers' training needs, you will probably have to sample probation officers.

In choosing a sampling design, you have a variety of alternatives. The major sample forms are listed below:

• **Nonprobability samples**	(Every unit does not have an equal chance for selection)
—Accidental	Take whatever units are available
—Quota	Take a certain number of each type of unit to fill quotas
—Purposive	Hand-pick the units to meet your needs
• **Probability samples**	(Every unit has an equal* chance for selection)
—Simple random	Select each unit at random
—Stratified random	Select a fixed percentage of each type of unit at random
—Systematic	Select every nth unit
—Cluster	Select a subgroup of the population and sample from that
—Combinations	Employ purposive and simple random samples, cluster and systematic samples, etc.

In outlining your sampling plan, you should define what is to be sampled (the inmate population, the parolees, the halfway houses). Some evaluations may have several populations to be sampled (volunteers, client records, counseling sessions).

*Actually, some sampling methods employ unequal selection probabilities that can then be corrected with inverse weights. The precise term to use here would be *a known nonzero probability* rather than *equal chance*.

In defining the populations, three characteristics should be specified:

- **Content** — What is the main area of interest—for example, skills, attitudes, characteristics?
- **Extent** — To what extent does the population vary in terms of geography, age range, profession, income, offense?
- **Time** — What is the time period being sampled—a point in time, such as December 31, 1975, or an interval, such as January 1, 1975, through December 31, 1975?

The population should be defined in terms of the characteristics or events to be sampled. For example, if one of the measures is *number of parolees*, a sample may be needed to obtain data on their characteristics, such as age, type of offense, release date, etc.

Next, you should select an appropriate design for each sample. If you are not going to select a probability sample, merely select as many members as you can of the population being sampled. If you are going to select a probability sample, you may need help. Any university and many consulting firms can supply it quickly and cheaply.

Finally, you need to select a sample size. Size will depend on sample design and the degree of precision you need. Again, you should consult an expert.

STEP 6: DEVELOP THE DATA COLLECTION PLAN

6.1 Review your needs (check the criteria and sampling standards)
6.2 Consider the alternatives: direct observation, direct reports, records
6.3 Outline the data collection plan (sources, methods, instruments, frequency, and timing).

The criteria and standards constitute lists of items that need to be collected; the sampling plan defines the sources. The timing specified in Phase I gives you the frequency and timing for collection of the data.

There are three principal data collection alternatives open to you: direct observation, direct reports, and records. The first two produce

what is sometimes called *primary data*. In corrections, there are many primary sources: inmate interviews, observations of training sessions, classification tests. Records are often called *secondary data*. The most common are statistical reports, special study reports, newspaper and journal articles, and evaluation reports. The following are the general data collection procedures available to you:

• **Direct Observation**	(The observer may be a participant or a nonparticipant.)
—Unstructured	No predetermined list of topics to be observed
—Structured	Specific topics selected for observation
• **Direct reports**	
—Free association	Respondent says anything he/she wishes; no set questions
—Unstructured interview	Selected topics are explored in depth
—Semi-structured interview	Open-ended questions
—Structured interview	Specific questions with a fixed choice of answers
—Tests	Psychological, medical, laboratory, etc.
—Inventories	Lists of characteristics, supplies, personnel, etc.
• **Records**	
—Statistical	Census, labor statistics, birth rates, etc.
—Documents	Legal, personal, government, business, etc.
—Secondary reports	Newspapers, research reports, etc.

In outlining your data collection plan, you should identify the data sources (the people, files, documents) and methods (observation, reports, records) that you will use. You should also identify the instruments that will be used (written tests, physical tests, face-to-face interview schedules). Finally, you should define the timing of the data collection: When will it begin? When will it end? How often will it take place?

STEP 7: DEVELOP THE REPORTING PLAN

7.1 Review your needs (audience, timing, frequency, purpose)
7.2 Consider the alternatives: oral, visual, both
7.3 Outline the reporting plan (medium, format, timing, style).

How will the data be reported—orally, in writing, or by a combination of these methods? What media will be used: graphs, slides, or memoranda? How often and when should the data be reported?

A report should be a statement of the findings on the effectiveness or efficiency of the subject evaluated.

To prepare your reporting plan, first review your needs. Who is the principal audience? What type of report would that audience prefer? Social scientists might prefer detailed written reports, while administrators, the public, and politicians might prefer short written or oral presentations. If there are multiple audiences, you may need multiple reports.

The purpose of the evaluation will guide you as well. The report should address the purposes identified in Phase I. If there are decision options, the report should include recommendations on those options.

The deadline for the evaluation and the frequency of evaluation will also guide you. Obviously, the report should be presented before the decision deadlines.

There are many ways to present a report. Among the more common oral and visual approaches are the following:

Oral	**Visual**
One-on-one	Article
Lecturer	Summary/abstract
Small group	Handout/throwaway
Panel	Slides
Conference	Flip charts
Mass media broadcast	Graphs
Question-and-answer	

Once the medium and format are chosen, the style of presentation should be selected. The style will depend greatly on the audience addressed and the media used. One group may prefer a technical presentation, another may prefer "plain talk," devoid of jargon. The

timing and frequency of reporting should also be determined. Will this be a one-shot report, or a periodic report? Finally, it is important to outline some procedures for feedback. The audience should be able to probe for clarifications as well as comment on the utility of the evaluation.

C. MANAGEMENT

STEP 8: DEVELOP THE MANAGEMENT PLAN

8.1 Review your needs (products, tasks, time-frame organization, resources)
8.2 Consider the alternatives: organization, resources
8.3 Outline the management plan (organization, staffing, tasking schedule, budget, monitoring, and supervision).

Once the technical plan is completed, you can put the management plan together. This plan should outline the tasks to be accomplished, the schedule for completing those tasks, the staffing and organizational pattern, the budget, and the monitoring and supervision strategy.

There are many ways to develop the plan. One is to identify your needs first, compare them with the resources available (or obtainable), and then reconcile the difference. For example, you can begin by identifying the products you need to develop (for example, an evaluation design, a pre-test report, a sampling plan, a final report). Then you can identify the tasks you need to carry out to develop those products (develop data collection instruments, draw a sample, interview parolees); the schedule that must be followed to complete the tasks on time; and the resources you will need to complete each task (staff, equipment, funds). You may also have some administrative requirements to consider, such as coordination among departments, control of data, access to inmates, etc.

There are some simple management tools that you can use.* One is a chart that includes the major tasks and products and the schedule.

*See Appendix B for samples.

Tasks and Schedule

Tasks	Jan	Feb	Mar	Apr	May	June
1. Complete design	—	—				
2. Complete instruments			*			
3. Draw sample		—	—			
4. Train interviewers			—			
5. Collect data				—	—	
6. Analyze data					—	
7. Prepare report						*

* = product

Another is a chart for estimating the personnel requirements for each task—broken down by type of skill and number of workdays, with the scheduled completion day for each task given.

Estimate of Persondays Required

Tasks	Day Due	Admin.	Eval.	Stat.	Sec.	Total
1. Complete design	15	2	7	1	1	11
2. Complete instruments	35		10	2	4	16
3. Draw sample	40		2	3	1	6
4. Train interviewers	40		2		2	4
5. Collect data	90		10		5	15
6. Analyze data	90	1	9	3	4	17
7. Prepare report	100	2	5	1	4	12
Total	100	5	45	10	21	81

Direct labor costs × daily rate	$94	$67	$60	$39	
Personnel	$471	$3015		$819	$4305
Consultant			$600		$ 600
Total labor					$4905

A third tool is a simple budget format you can also use to estimate other resource requirements and the total estimated cost of the evaluation.

Project Resource Requirements

Item	$	Description
Personnel	4,305	Administrator, in-house evaluator secretary (71 days total)
Fringe benefits (10%)	431	
Consultants	600	Statistical consultant (10 days)
Supplies/ materials	100	Office supplies, interview forms
Equipment	—	No additional equipment needed
Facilities/space	—	No additional equipment needed
Postage/freight	—	None
Reproduction	50	Copies of interview instruments
Telephone	50	Calls to consultant
Travel	80	Consultant travel
Per diem	40	Consultant per diem
Miscellaneous	100	Contingency
Total	5,756	Total direct costs, no overhead charges

When it comes to resources, the principal ones to consider are funds and staff. You probably already know how much money you have in your regular budget for evaluation. If additional funds are required, you may get them from discretionary budgets, special funds, grants, contracts, and so forth.

Staff can come from inside or outside your agency. Stuart Adams identified four alternatives for staffing evaluation efforts.* The advantages and disadvantages of each are summarized here.

Op. cit., pp. 29-33. Also see Weidman, *op. cit.*, pp. 33-39.

In-House Research Staff

Advantages	Disadvantages
Usually does best work	May lose objectivity
Staff familiar with program	Difficult to attract and retain qualified staff
Readily available and accessible	Costly to maintain permanent staff
Follow-up of evaluation possible	

University Faculty

Advantages	Disadvantages
Highly trained and skilled	May be inflexible, highly theoretical, impractical
Readily available	Limited time available
	Lack of interest
	Use of incomprehensible jargon

Private For-Profit Research Firms

Advantages	Disadvantages
Strong in technique and methodology	Weak in correctional knowledge and theory
Interested in client's programs	Naive about agency's objectives and procedures
Businesslike	Some overly concerned about making profit
Concern for doing work that the client will judge satisfactory	Often use incomprehensible jargon

Private Non-Profit Research Firms

Advantages	Disadvantages
Usually strong in correctional knowledge and experience	Evaluation staff not usually specialists
Committed to advancing the field	Limited number of firms
Willing to do pioneering work	Usually small in size

Other sources of personnel include State Planning Agencies, volunteers, college students looking for thesis topics, and corrections

staff and clients. Some researchers have suggested using guards and inmates to collect data.

Once you have considered the alternatives, you should be ready to outline your management plan:

1. Specify the organizational structure, staffing, roles, and relationships. This involves assigning tasks to individuals, clarifying decision authority and reporting relationships, and describing what each individual will do and whom he/she will work with.
2. Complete the tasking and scheduling. This can be done by revising the chart you have already developed.
3. Complete your budget. Again, you can simply revise your draft budget.
4. Outline your monitoring and supervisory system. The organizational structure, staffing pattern, tasking, schedule, and budget may be enough for a simple management information system. You can use these plans to monitor the performance of the evaluation.

IV. PHASE III: CONDUCTING AND MANAGING THE EVALUATION

The third and last phase of an evaluation process is the implementation of the evaluation design and the production of the results. Phase III consists of five steps: 1) making staff assignments for the evaluation; 2) developing the evaluation and management procedures; 3) pre-testing and revising the evaluation procedures; 4) collecting and analyzing the data and reporting the results; and 5) developing strategies for using the evaluation findings. Ideally, the product of this phase will be information that fits precisely the purpose that you identified in Phase I.

STEP 1: MAKE STAFF ASSIGNMENTS FOR THE EVALUATION

Once you have the go-ahead on your proposed methodology, you can begin. Formalizing the staff assignments is often a first step. Since you have already developed your organizational plan, staffing pattern, and so forth, you should be ready to go. If you will be using new staff, you should allow time for their recruitment, selection, and training.

STEP 2: DEVELOP THE EVALUATION AND MANAGEMENT PROCEDURES

2.1 Review your needs (from Phase II: measurement, sampling, data collection, data analysis, reporting, management procedures)
2.2 Consider the alternatives: develop them yourself, consult experts
2.3 Develop the procedures.

Unless you have a very simple evaluation methodology, you will probably have to develop some data collection instruments, computer processing instruments, and so forth. There are many standard texts that deal with these subjects and quite a few experts whom you can consult if you need help.

Here are some of the procedures you may need to develop or have developed:

- **Measurement** — Operational definitions of the measure (how to define contacts, visits, direct costs, recidivism); construction of measures (averages of inmate days, measures of skill gain); instructions (how to derive dropout rates from other measures)
- **Sampling** — Operational definitions of the populations to be sampled (how to define the inmate population, the halfway houses to be sampled); procedures for calculating sample sizes, sampling error, confidence limits; procedures for drawing the sample (list the names of all participants, assign numbers, select a number from a random number table)
- **Data collection** — Construction of the instruments (interview schedules, tests, questionnaires); instructions to interviewers; training procedures for interviewers; forms for recording data

- **Data analysis** — Instructions for classifying data (group the inmates by five-year age groups); coding instructions (code "Yes" as 1, "No" as 2); tabulation instructions (cross-tab variables 2 and 5); statistical procedures (regression, chi-square). If you are planning to use computers, you may also have to develop programs, as well as key punch, editing, sorting, or other processing instructions
- **Data reporting** — Report forms, reporting instructions, printout formats
- **Management** — Detailed budgets, schedules, job descriptions, personnel records, financial reports, staff communication procedures.

Once you have listed your needs, you will be in a better position to determine whether you and your staff can develop the procedures or you will need outside help. This may be an appropriate time to review your tasking and scheduling. Now that you have a more detailed view of the technical and managerial needs, you may want to revise your management plan to provide yourself with the level and type of assistance you need.

STEP 3: PRE-TEST AND REVISE THE EVALUATION PROCEDURES

3.1 Review your needs (time, resources, staff, methodology)
3.2 Consider the alternatives: review, simulation, field tests
3.3 Implement the pre-test (select the procedure, conduct the test, refine the evaluation).

If the evaluation procedure is simple, you may not need to pre-test. To decide whether and what type of test you need, consider four factors:

- **Time and resources available** — Where you have little time and few resources for evaluation, you would probably settle for limited procedures.

- **Complexity of the procedures** — The greater the complexity and sophistication of the evaluation procedures, the greater the need for a careful test, so that you can work out any problems before the field work begins.
- **Significance of the evaluation** — The more important evaluations should be carefully tested to make sure that the information desired will be produced.
- **Skills of the staff** — If you have any doubts about the skills of your staff, or if you have a newly trained staff, a pre-test is a good opportunity to give them some practical experience.

Once you have identified your needs, you should select an appropriate testing procedure. Three common approaches are:

- **Review** — Involve experts, typical subjects, or some of the target audience in examining the evaluation design and procedures. Do they appear to them to be on target, understandable, practical?
- **Simulation** — This could be done by role playing or by manufacturing hypothetical data to test procedures.
- **Field test** — This would involve a pilot run of the data collection, analysis, and reporting procedures, conducted at a site similar to the one where the evaluation will be implemented.

Once you have selected your approach and conducted the test, the results can be used to revise the topic, the methodology, the procedures, or any other part of the evaluation.

STEP 4: COLLECT AND ANALYZE THE DATA, AND REPORT THE RESULTS

4.1 Implement the evaluation plan (draw the sample, collect the data, analyze the data, report the findings)
4.2 Identify and solve problems.

You are now ready to implement your evaluation procedures. Implementation involves:

- ***Drawing the sample***
- ***Collecting the data***
- ***Analyzing the data***
- ***Reporting the results.***

Your principal concerns at this point are to follow the evaluation plan and procedures that you have refined and to solve any problems that might arise. Here are some typical problems that have been encountered in corrections evaluation:*

- **Interaction of Evaluation and Program Operations**

Sometimes the evaluation interferes with program operations—for example, when an inmate survey causes delays in work schedules or leads to grumbling and unrest. At other times it is the reverse—that is, the program operations interfere with the evaluation—such as when inmate work schedules interfere with interviews.

- **Discrepancies Between Planned and Actual Program Activities**

Sometimes the subject to be evaluated does not operate as planned. For example, the counseling services are not provided; the recreational schedules are not followed; the planned probation visits do not take place; the grievance procedures are changed; or the training program is delayed. A common problem in controlled experiments is contamination, where a control group that is not supposed to get an experimental program is allowed to participate because "it would be unfair to deny them the opportunity."

- **Changes in the Needs of Administrators**

Sometimes an administrator changes his/her mind about an evaluation, either because of second thoughts, a shift in priorities, or a declining need for the information. Organizational and personnel changes can also affect the evaluation—for example, when the warden who wanted the evaluation is replaced by one who has other needs and interests.

- **Technical Difficulties**

Finally, there are the problems of unavailable data, incomplete samples, errors in data collection, bugs in the computer program, coding problems, and so forth.

*For a more detailed discussion, see Weidman, *op. cit.*, pp. 13-16, 26-30.

Typical solutions to these problems usually involve serious management decisions:

- **Intervention by the Evaluator**

The evaluator might step in to straighten out a program that is not operating as planned or advise the program staff on how to perform so that the evaluation is not jeopardized. To do this successfully, the evaluator's authority for program operations must be clearly specified and understood by all.

- **Intervention by the Administrator**

The administrator might assume this role, taking a firm hand in conducting the operation of the subject being evaluated.

- **Modification of the Evaluation Design**

This could involve one or several parts of the design—modification of the data collection procedures, alternation of the study design, changes in the timetable, and so forth. Usually, a change in one part of the design will have implications for all other parts.

- **Termination of the Evaluation**

This is a radical step, but if the evaluation is no longer able to produce valid and reliable information, it may be prudent to halt, and re-invest the resources in something more productive.

- **Continuation with No Changes**

This can be equally radical, if the evaluation has been compromised significantly. But the administrator and evaluator may feel that some information will be better than none and decide to continue.

STEP 5: DEVELOP STRATEGIES FOR USING THE EVALUATION FINDINGS

5.1 Review the evaluation findings (primary report, other literature)

5.2 Evaluate the evaluation (did it meet its objectives, was it efficient?)

5.3 Develop action strategies (take no action, disseminate results, make recommendations, develop new programs, modify existing programs, terminate programs, investigate further).

Evaluation findings should be reviewed from the perspective of the decision options (or evaluation purposes) identified in Phase I.

In reviewing the evaluation report, both the evaluator and the administrator should anticipate that they might encounter some unexpected or undesirable findings. These must be dealt with alongside the expected and desired findings. Also, other literature should be reviewed as a matter of routine to supplement the findings of the primary reports. In fact, evaluators should set up information accumulation and screening systems that would provide administrators with summaries of relevant research and evaluation findings.

You should also evaluate the evaluation. Both the evaluator and the administrator should assess the utility of the evaluation. Did it meet the objective stated? Was it conducted efficiently—that is, on time and within the budget planned? Are the results useful? Adams noted that "Researchers and funders of research in corrections often complain that research products are not used by correctional administrators and their staffs."* There are some good reasons for this:

- The recommendations are not acceptable to the administrator.
- The staff does not agree with the findings.
- The criteria and standards are unacceptable.
- The subject matter is irrelevant.
- The findings are not useful.
- The recommendations are not feasible.
- The findings are inconclusive.
- The findings are incomprehensible.

Most of these constraints to use of the evaluation can be overcome by following the evaluation guidelines that have been presented. The problems reflect a lack of communication between the evaluator and the user. If the target audience is identified and involved in specifying the topic, developing the evaluation plan, and implementing the evaluation, there is every reason to believe that the results will be useful. But even that cannot be left to chance. Once the results are in and reviewed, you must evaluate the evaluation to determine whether any changes in its topic, design, or conduct must be made in the future. Finally, you need to outline some action strategies for using the findings.

You should begin by relating the findings to the original purpose of the evaluation. Now that you have the data, are the options still

*Weidman, *op. cit.*, p. 34. See also pp. 31-32.

workable? Should they be modified or dropped? Should new options be added?

Next, you can identify the courses of action open to you and some strategies for using the evaluation results:

OPTIONS	*STRATEGIES*
• **Disseminate the findings**	Disseminate to the target audience as planned
	Send summaries to key figures in the legislature, SPA, LEAA, etc.
	Telephone the press
	Present to a professional journal or convention
	Incorporate into agency literature
	Integrate into staff training courses
	Distribute reports to similar agencies
• **Recommend action**	Make recommendations to the target audience as planned
	Identify other key actors who might take action and make recommendations—the governor, the parole board
	Recommend action to a constituency or interest group, such as the ACA, community groups
• **Take action**	Review the action options identified and select one as planned
	Feed the results into planning by involving the evaluator in planning sessions
	Feed the results into program development by organizing an advisory group or task force to recommend a new program approach
	Propose budget increases using the evaluation as backup; take the evaluator to the budget hearings
	Modify the subject evaluated

Terminate the program

Begin action to evaluate the subject in more detail.

With the selection of decision options and action strategies, the evaluation process completes its cycle by feeding information and judgments into management to help administrators make decisions about program planning and operations. If the results are useful, one of those decisions may be to conduct another evaluation—and you begin again with Phase I.

APPENDIX A

SUMMARY OF THE STEPS IN THE EVALUATION PROCESS

PHASE I: SELECTING THE EVALUATION TOPIC

STEP 1: Identify the Evaluation Priorities

1.1 Identify your evaluation interests—consider and list questions to be answered, hypotheses to be tested, judgments and decisions to be made, program goals and objectives, problem areas, major issues

1.2 Rank order these interests according to their utility—consider utility for management, public relations, accountability, reputation

1.3 Negotiate consensus on the order of ranking (priorities)

STEP 2: Clarify the Evaluation Subject

2.1 Review the relevant literature, documents, records

2.2 Identify the object, person, or event to be evaluated

2.3 Determine the number of subjects to be evaluated

2.4 Identify any features to be given special emphasis

2.5 Identify the type of evaluation to be conducted: need, design, performance, impact

STEP 3: Clarify the Objectives of the Subject

STEP 4: Specify the Objective of the Evaluation

4.1 Identify the target audience of the evaluation

4.2 Define the purpose of the evaluation in relation to the target audience: information, judgment, decisionmaking

4.3 Specify the decisions and the decision options the evaluation should address

4.4 Determine the deadline(s) for the evaluation and the frequency of conducting it

STEP 5: Ascertain the Feasibility of the Evaluation

5.1 Determine whether the evaluation can be conducted

- Identify any technical problems that might make the evaluation impossible: lack of data, no objectives, no way to measure impact
- Estimate the resources potentially available and compare those with what might be required; or determine what could be done with the available resources of staff, money, time, and equipment

5.2 Determine the probability that the results will be used

- Identify the internal and external constraints that might inhibit the use of the evaluation results

- Identify the positive internal and external factors that might support use of the evaluation results

5.3 Revise the subject or objective, if necessary

PHASE II: DEVELOPING THE EVALUATION PLAN

A. THE JUDGMENT STEPS

STEP 1: Select the Study Design

1.1 Review your evaluation needs (check the purpose, audience, timing, constraints)
1.2 Consider the alternatives: experimental vs. non-experimental designs
1.3 Select a design

STEP 2: Select the Evaluation Criteria

2.1 Review your needs (check purpose, decision options, program objectives, emphasis, and constraints)
2.2 Consider the alternatives: effectiveness or efficiency or both
2.3 Select the criteria
2.4 Assign weights (unweighted, ranked, scores)
2.5 Select the measures (counts, rates; ratios, proportions, percentages, indexes, formulas)

STEP 3: Select the Evaluation Standards

3.1 Review your needs (see the type of evaluation and study design)
3.2 Consider the alternatives: need, demand, plan, past performance, similar subjects, control groups, requirements, professional standards, optimal model
3.3 Select the standards
3.4 Assign weights
3.5 Select the measures

STEP 4: Develop the Analysis Plan

4.1 Review your needs (review the purpose, study design, criteria, and standards)
4.2 Consider the alternatives: statistical or nonstatistical
4.3 Outline the plan: classification, codification, tabulation, statistical manipulation, interpretation

B. THE DATA PROCESSING STEPS

STEP 5: Develop the Sampling Plan

5.1 Review your needs (check the study design and criteria)

5.2 Consider the alternatives: nonprobability vs. probability samples
5.3 Outline the sampling plan: population, sample design, sample size

STEP 6: Develop the Data Collection Plan

6.1 Review your needs (check the criteria and sampling standards)
6.2 Consider the alternatives: direct observation, direct reports, records
6.3 Outline the data collection plan (sources, methods, instruments, frequency, and timing)

STEP 7: Develop the Reporting Plan

7.1 Review your needs (audience, timing, frequency, purpose)
7.2 Consider the alternatives: oral, visual, both
7.3 Outline the reporting plan (medium, format, timing, style)

C. MANAGEMENT

STEP 8: Develop the Management Plan

8.1 Review your needs (products, tasks, time-frame organization, resources)
8.2 Consider the alternatives: organization, resources
8.3 Outline the management plan (organization, staffing, tasking schedule, budget, monitoring, and supervision)

PHASE III: CONDUCTING AND MANAGING THE EVALUATION

STEP 1: Make Staff Assignments for the Evaluation

STEP 2: Develop the Evaluation and Management Procedures

2.1 Review your needs (from Phase II: measurement, sampling, data collection, data analysis, reporting, management procedures)
2.2 Consider the alternatives: develop them yourself, consult experts
2.3 Develop the procedures

STEP 3: Pre-test and Revise the Evaluation Procedures

3.1 Review your needs (time, resources, staff, methodology)
3.2 Consider the alternatives: review, simulation, field tests
3.3 Implement the pre-test (select the procedure, conduct the test, refine the evaluation)

STEP 4: Collect and Analyze the Data, and Report the Results

4.1 Implement the Evaluation Plan (draw the sample, collect the data, analyze the data, report the findings)

4.2 Identify and solve problems

STEP 5: Develop Strategies for Using the Evaluation Findings

5.1 Review the evaluation findings (primary report, other literature)

5.2 Evaluate the evaluation (Did it meet its objectives? Was it efficient?)

5.3 Develop action strategies (take no action, disseminate results, make recommendations, develop new programs, modify existing programs, terminate programs, investigate further)

APPENDIX B

WORKSHEETS

B-1
WORKSHEET FOR SUMMARIZING THE EVALUATION PLAN

PHASE I: SELECTING THE EVALUATION TOPIC

1. Priorities _____
2. Subject and Type of Evaluation _____
3. Subject Objectives _____
4. Evaluation Objective _____
   - Target Audience _____
   - Purpose _____
   - Decisions/Options _____
   - Deadline/Frequency _____
5. Feasibility _____

PHASE II: DEVELOPING THE EVALUATION PLAN

1. Study Design _____
2. Criteria _____
3. Standards _____
4. Analysis Plan _____
5. Sampling Plan _____
6. Data Collection Plan _____
7. Reporting Plan _____
8. Management Plan _____
   - Tasks _____
   - Schedule _____
   - Cost _____

PHASE III: CONDUCTING AND MANAGING THE EVALUATION

1. Evaluation Staff _____
2. Evaluation and Management Procedures _____
3. Pre-test and Revision _____
4. Collection, Analysis, and Reporting _____
5. Use of Findings _____
   - Evaluation of the Evaluation _____

B-2
WORKSHEET FOR PHASE I: SELECTING THE EVALUATION TOPIC

SUMMARY: To provide _____with evaluative information on _____

by _____in order to _____

PRIORITY NO.

SUBJECT	AREAS OF EMPHASIS	TYPE EVAL.	OBJECTIVES

	DECISIONS	OPTIONS	TARGET AUDIENCE	DUE DATE

B-3
WORKSHEET FOR PHASE II:
DEVELOPING THE EVALUATION PLAN

CRITERIA	MEASURES	SAMPLE/DATA

SOURCE	DATA	ANALYSIS

B-4
TASKS AND SCHEDULE WORKSHEET

DAYS/WEEKS/MONTHS

TIME LINE	
TIME / TASKS	

B-5
TASKS AND STAFFING WORKSHEET

TASKS	DUE BY DAY	ESTIMATE OF PERSONDAYS/
TOTALS		

DIRECT LABOR COSTS

$ COST X TIME =

HOURS REQUIRED

	TOTALS

B-6
BUDGET WORKSHEET

ITEM	$
Direct Labor Costs	
Fringe Benefits	
Consultants	
Supplies/materials	
Equipment	
Facilities/space	
Postage	
Reproduction	
Telephone	
Travel	
Per diem	
Miscellaneous	
Other	
Total	

	DESCRIPTION

END