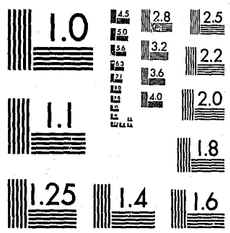

ARRO-3039-TR

NATIONAL EVALUATION PROGRAM PHASE I
ASSESSMENT OF CORRECTIONAL PERSONNEL TRAINING PROGRAMS
VOLUME 3. EVALUATION ISSUES AND STRATEGIES

Howard C. Olson, Project Director
Merri-Ann Cooper
Albert S. Glickman
Robert Johnson
Shelley J. Price
Ronald I. Weiner

Technical Report

Prepared under contract to the Office of Program Evaluation, National Institute for Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration

Contract No. J-LEAA-023-78

This project is being supported by Contract #J-LEAA-023-78 awarded by the 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.

Advanced Research Resources Organization
March 1980

NATIONAL EVALUATION PROGRAM PHASE I
ASSESSMENT OF CORRECTIONAL PERSONNEL TRAINING PROGRAMS
VOLUME 3. EVALUATION ISSUES AND STRATEGIES

Howard C. Olson
Merri-Ann Cooper
Albert S. Glickman
Robert Johnson
Shelley J. Price
Ronald I. Weiner

Technical Report

March 1980

R80-2

TABLE OF CONTENTS TO FULL REPORT

TABLE OF CONTENTS--VOLUME 3

LIST OF TABLES

LIST OF FIGURES

PREFACE

This Phase I report deals with the state-of-the-art of training of correctional personnel. On the basis of the knowledge gained by way of visits to 17 correction agencies across the nation, a questionnaire mailed to a national sample of corrections agencies, plus review of literature dealing with the field of program evaluation, evaluation models appropriate to correctional personnel training have been generated and documented. The models are intended to be general techniques that an agency may apply to its own training programs without the aid of additional evaluation experts. The Phase II plan for demonstrating the models is outlined.

The report is organized as 12 chapters and published in 4 volumes:

Volume	Chapters	Title
1	I	Executive Summary
2	II - IX	Correctional Personnel Training: Conceptual and Empirical Issues
3	X - XII	Evaluation Issues and Strategies
4	-	Appendixes

ACKNOWLEDGEMENTS

Many persons other than the ARRO project staff have contributed significantly to the project. Among others, Mr. Thomas Rosazza of the Maryland Training Academy, Lieutenant Jeffrey Paskow of the Montgomery County (MD) Training Academy, and Corporal Robert Strickland of the Fairfax County (VA) Jail gave of their time and expertise to assist in the development and pretest of the Interview Guides and questionnaire used in the national survey of correctional personnel training.

Our hosts at the various training sites visited were outstanding in their cooperation, reception of team members, and support for the project. Training directors and their staffs spent much time sharing with us their experience with and knowledge of the training process.

Those who responded in the national survey demonstrated genuine concern and conscientious efforts to provide the information requested. Without the interest and assistance of the many individuals involved in the training of correctional personnel, the goals of the first phase of this project could not have been accomplished.

The government project monitors, Dr. Harold Holzman, succeeded in December 1979 by Dr. Richard Layman, have created the climate for an excellent working relationship with the Law Enforcement Assistance Administration. They have been understanding and supportive, providing valuable comments and guidance throughout.

Several former members of the ARRO staff contributed significantly to the project. Dr. Jeffrey Kane constructed the original conceptual models of a generic training system. Ms. Nancy Yedlin assisted in project management, participated in many site visits, and helped draw out site visit findings. Ms. Sharyn Mallamad organized site visits and developed initial versions of the evaluation strategy proposed. Special recognition is due Ms. Patti Vernacchio for her outstanding secretarial talents, organizational abilities, and high enthusiasm and drive throughout the project.

CHAPTER X

ISSUES IN EVALUATION

INTRODUCTION

One of the final goals of this project is to develop an evaluation model that can be used by correctional training personnel to assess and improve their training efforts. Unfortunately, such a task is not as simple as handing over a prototype evaluation model and expecting it to be workable, understood, and applied. The conduct of an evaluation calls for much more than the mechanical application of models and methodologies suggested by experts. While a workable model for corrections needs to be general and widely applicable; in practice, evaluation must be guided by a situation-specific strategy. The user needs to select parts from the model and adapt it to individual and organizational surroundings. Selection of the proper strategy and evaluation implementation require an awareness of affective conditions in corrections as they relate to the training process, as well as a sensitivity to issues in evaluation. In Chapter IV, issues of concern in correctional training are presented in order to familiarize the practitioner with some of the situational variables we view as having impact upon the training and evaluation process. In order to provide a basic knowledge of the literature and a picture of training evaluation activities in corrections, this chapter contains a discussion of matters directly related to evaluation alone.

Several basic questions about evaluation are addressed. First, why should we bother to evaluate? Second, despite the deluge of literature and professional advocacy for training program evaluation, there have been few instances of rigorous evaluations of correctional personnel training programs. What are some of the factors contributing to this lack of evaluation activities? Third, what are some characteristics of the evaluation that have been conducted? Finally, processes used to identify job responsibilities and needs in corrections will be described. This chapter is intended to clarify some assumptions evident in the design of the refined evaluation model and will lay the groundwork for its presentation in the chapter following.

WHY EVALUATE?

It is difficult to convince agency administrators and training staff of the merits of evaluation when training units are characteristically understaffed, overloaded, and short on funds. Despite the skepticism with which evaluation is received, there are nonetheless substantial reasons why one should expend time and effort to evaluate training efforts--to aid in decision making, to satisfy accreditation standards, to provide information about individual and organizational needs, to establish cost-effectiveness, to demonstrate competence in the organization.

Facilitate Decision Making

Information obtained through evaluation can facilitate decision making by reducing uncertainty (Patton, 1978). Imagine a training director who anticipates a budget cut for the upcoming fiscal year. The director has decided to absorb the loss of funding by eliminating a portion of the existing training program. Having a firm idea of which course should be discontinued, the director choses to conduct an evaluation to verify his choice. An evaluation designed to address questions such as the following may provide information for this purpose.

- Is there still a need for the course?
- Do trainees find the course interesting and relevant to their job?
- Have trainees learned the information presented in the course?
- Do trainees use this training back on the job, i.e., has their performance changed?
- Has the course affected organizational functioning.

To be worthwhile, an evaluation does not always have to uncover unexpected or earth-shattering findings; sometimes it may be valuable because it confirms what training staff have suspected all along. Evaluation results can serve to substantiate hunches, as well as provide hard evidence as impetus for change (Patton, 1978).

Provide Information About the Needs of the Trainees, the Organization, and the Job

Quite frequently, a training program is established without consideration of characteristics of the trainee population and requirements of the job they are to perform. Needs of the trainee's home organization may similarly be overlooked in developing a training program. Education and training in all sectors, as well as in corrections, can easily fall prey to faddish instructional techniques and gimmicky programs. The lure of promising advertisements for innovative and easy programs is sometimes attractive enough that an organization will invest in a flashy program before determining whether or not such training will facilitate job performance or relieve organizational difficulties. Training programs aimed at developing behavior modification skills or inmate counseling techniques are valuable in many instances (Grzegorek & Kagan, 1974, Hosford, George, Moss, & Urban, 1975; Smith, Milan, Wood & McKee, 1976). However, in situations where the trainee population consists of semi-literate employees, or the organization is experiencing a barrage of lawsuits, or maintenance and security tasks are carelessly performed, rudimentary and remedial training in procedures and report writing may be more appropriate than more advanced and novel programs.

In the current economic crunch, many applicants for line level jobs in corrections have advanced college degrees. Yet EEOC requirements necessitate the hiring and training of staff with less educational background. When training staff are called upon to work with such a wide range of individuals, an awareness of needs and abilities is critical. Training development and evaluation techniques are tools to help the trainer identify needs and design a training program accordingly.

Accreditation and Mandates

In corrections, training has become generally a mandated activity. Evaluation of training has not achieved the same status. Yet, many government and private agencies strongly recommend the inclusion of program evaluation in training plans. In 1967, the Joint Commission on Correctional Manpower and Training declared that "evaluation is a necessary

part of continued program development" (p. 3). Practice standards issued by the Commission on Accreditation for Corrections stipulate that training and staff development programs should be formally evaluated at least once a year (1977c). Frequently, agencies that receive federal funding to operate training programs are required to conduct impact evaluations. The Law Enforcement Assistance Reform Act (LEARA, 1979, The Bellringer, 1979) recently passed by the Senate, (and anticipated to pass the House without major changes), includes amendments that have serious implications for the evaluators of LEAA funded criminal justice programs. The Biden amendments contain provisions that place emphasis on evaluation of programs/projects, including those receiving funding for the training of criminal justice personnel. The act stipulates that grant applications to LEAA must also include detailed provisions for project evaluations. Program evaluation can be an asset, if not a requisite, when applying for grants, funding, and accreditation.

Establish and Maintain Cost-Effectiveness

"Properly performed course evaluations not only earn their own way by leading to improvement in course quality, but in a great many cases, reveal opportunities to bring about substantial reductions in training costs" (Lott, 1967, p. 38). Consider, for example, the classic problem in corrections training--the difficulty of releasing persons from the job to participate in training. At times, the training staff may be forced to hold classes with very few participants. In effect, general budget priorities that place low value on relief staff fail to take into account the loss of money spent on empty seats in small training classes. Evaluative data documenting such cost inefficiencies can make a case for better coordination of programming and employee releases. Perhaps, no better way exists for training personnel to win the attention of administrative officials than by suggesting ways to save money.

Aside from organizational and departmental desires to make the cost of a tight budget, there is a need to convince those agencies providing funding that their money is being handled wisely. Human service programs

are increasingly under pressure to demonstrate accountability in terms of cost. The evaluation process can aid program staff in achieving and demonstrating cost-effectiveness.

Demonstrate Competence

When the training department is at the bottom of the budget priority list, training staff are constantly required to prove their worth (Alden, 1978), and trainers have increased legal liability for violations of inmate rights, evaluation can be especially useful. If training staff initiate evaluation activities, they will have more choice about how the evaluation is conducted--what questions are asked, what methods are used, how data are collected, and so on. Ongoing evaluation, enabling the trainers to continually monitor and revise the program, demonstrates interest in doing the best job possible. By staying a step ahead of those who might challenge the validity of the training program, the training unit has a better chance of maintaining a position of stability within the organization.

PROBLEMS IN EVALUATION

Despite insistent declarations of worth and repeated demands for program evaluation, correctional administrators and training staff often exhibit a strong resistance to evaluation for reasons that go beyond insufficient time and funding. Part of the skepticism of evaluation comes from a distrust of social scientists who fluently critique the correctional system for its lack of sophistication, yet, according to practitioners, are unaware of true dilemmas beyond the theoretical level. Many workers in the field have seen tremendous sums spent on research and development activities that have given no relief from existing problems such as institutional overcrowding or oversized caseloads, proposing either no specific technique to ameliorate problems or offering ones that cannot be implemented. Yet, the issue is not solely a wary distaste for professional intervention and advice. Some very real concerns related to the evaluation process itself contribute to the widespread hesitance to accept and apply evaluation procedures. Many reservations reflect a fear that evaluation may fail to produce the expected

impact, (i.e., act as impetus for needed change) or information acquired through an evaluation may be misunderstood and used inappropriately. Training personnel with whom we spoke described themselves in a double-bind. If an evaluation demonstrates program effectiveness, who cares? If the evaluation reveals problems, the program is killed. The training director is in a no-win situation. The sections following discuss external factors that influence decisions about whether or not to evaluate.

Decision Making, Politics, and People

Despite an earlier assertion that information obtained through evaluation can aid the administrator or policymaker in making program decisions, this often is not the case. Unfortunately, for a variety of reasons, the decision making process may not include consideration of evaluation findings. Frequently, the evaluator has not addressed or has been forced off from the questions regarded as relevant by the decision maker. Political preferences may not agree with findings, and policy decisions are greatly influenced by public sentiment. The "pork barrel" evaluation model, as described by Patton (1978), provides an example. The criterion of success when this model is applied is the political clout of program constituents. If those who favor the program are in positions of power or influence, regardless of true program value, the reviews will be positive.

Political overtones aside, evaluation is dependent on individual efforts. Whether or not an evaluation is carefully conducted, and in turn leads to change, depends primarily upon whether or not someone is motivated enough to head up the effort and see it through. An evaluation of statewide correctional training efforts over a 4-year period revealed that until a training director was hired, virtually nothing new was undertaken by training staff. With the advent of a director, changes (including evaluation) began to occur (Robinson, 1975). The leadership itself may be the most important factor in program implementation and evaluation (Anderson & Ball, 1978; Patton, 1978).

Results and Repercussions

Evaluations often are viewed by practitioners as self-defeating. The more carefully conducted and controlled the evaluation, the greater is the likelihood of results that show deficiencies in the training program (Bunker & Cohen, 1978; Newstrom, 1978). Unfortunately, evaluation information that indicates room for improvement (as there invariably is) often is interpreted to mean a program is "bad." The resulting label may lead to a layoff of staff members, funding reductions, or elimination of the training program altogether (Hanford & Moone-Jochums, 1978). No wonder that program personnel may feel that when embarking upon evaluation they are putting themselves on report.

Confusion

Once the training staff decides to undertake a program evaluation, the question becomes, "Where do we start?" Recipes for evaluations are plentiful. It becomes difficult for anyone except professional evaluators to distinguish between faddish techniques and tried-and-true evaluation strategies. There is the Adversary Model, the Medical Model, Formative Evaluation, Summative Evaluation, Transactional Evaluation, Discrepancy Evaluation, Goal-Free Evaluation (Anderson, Ball, Murphy, & Associates, 1975)--each endorsing a different design and data collection method. The corrections practitioner is faced with confusing and even contradictory advice (Patton, 1978). Controversy arises concerning "hard" vs. "soft" methods of evaluation; an advocate of the former assumes the best (and only) way to evaluate is to apply the most rigorous design and sophisticated quantitative analysis methods. Those who promote less rigid and more qualitative evaluation strategies claim that statistics produced by "hard" evaluations ignore sensitive and important program aspects. Such theoretical debates, where adherents of different views argue the legitimacy of different approaches, create a "no-win" atmosphere as the corrections practitioners experience it, clouded by infeasibility and high cost.

Evaluators do not purposely devise new strategies to confuse and alienate consumers of their trade and tools. There are many evaluation strategies because there are many different types of programs that vary in purpose, structure, content, clients, and resources. The competent evaluator tries to be alert to evaluation complexities and the political implications noted earlier.

TRAINING EVALUATION IN CORRECTIONS

Despite the gloomy outlook when comparing evaluation in theory and evaluation in practice, all correctional training staff members are not rigid in their beliefs nor stubborn adherents of the status quo. Some realize the benefits of evaluation and overcome numerous difficulties to implement serious evaluation efforts. Not until recently, however, since correctional programs have increased in visibility and need for accountability, have evaluations been included in program plans or accorded sufficient funding.

When evaluations that have been conducted are reviewed, several characteristics occur with predictable regularity. Most evaluations in corrections are not written into program plans and are thus conducted ex post facto--virtually eliminating the possibility of controlling experimentally or statistically many of the variables that influence trainee learning and performance. A majority are conducted by outside consultants, the most costly and transitory kind of evaluation (transitory in that the evaluation is not continual, providing feedback for ongoing program improvement, and thus likely to be weak and temporary in its effects). Evaluation too often is a one-shot, post hoc deal, pertaining only to the group of trainees studied. Rarely, is the training program and its evaluation designed simultaneously beforehand, as they should be. There is generally failure to link training with on-the-job performance; to assess transfer of learning and skills from the classroom to the job. Most evaluation results point to a need for the establishment of clearly defined organizational and training program goals and objectives, corollary inspection of the skills and abilities required to perform the job, and continuing monitoring and feedback of

the implementation of recommendations. Assessment of how interesting and how enjoyable the training has been is far more common than attempts to establish whether or not the training has fulfilled on-the-job needs (if, indeed, needs have even been identified).

Some measures of correctional training program effects, extracted from evaluation reports, are described in the section following. Concerns and constraints associated with each type of measure will be discussed.

Attitudes and Opinions about a Program

Self-report questionnaires are the most common method for evaluating training. Trainees generally are asked to rate aspects of the program--quality of instruction, adequacy of facilities, course presentation, relevance of training content, suitability of training materials, length of training sessions, fairness of tests, and so on. Most evaluation forms provide rating scales with various program aspects assessed on a 4- or 5-point scale. The ratings may be supplemented by open-ended questions about the "most valuable" or "least valuable" part of the training and what might be added or omitted to improve the course. General comments are invited. Sometimes a brief evaluation form is distributed following presentation of each topic or module of the training session (Site Visit Observations [SVO]; Harris & Harris, 1972).

Another evaluation tactic asks questions concerning the extent to which expectations of training participants have been met. At the beginning of the session, trainees are requested to make a list of what they would like, or expect to get out of training. At the close of the program, the participants refer back to these objectives to assess the degree to which their desires or expectations have been satisfied. More frequently this occurs only as a post hoc retrospective exercise.

Finally, trainers have relied upon the reactions of an objective observer in order to obtain evaluative feedback. In this instance, an observer attends the training session and keeps record of events and reactions, and either prepares a subjective narrative (West Virginia

University, 1970), or completes in greater detail a set of ratings and/or responses to structured questions bearing upon various aspects of the training.

It would be fair to say that most of these evaluations and rating forms provide, primarily, information about the personality and style of the trainers or enjoyability of the subject matter. Although it is desirable that programs score well on these indices, seldom are strong relationships shown between responses on opinion questionnaires and job performance. Most often no attempt is made to establish the relationship.

Pretests and Posttests of Learning and Attitude Change

Objective tests designed specifically from course content are sometimes given prior to training and again following the course. Composed of multiple choice, fill-in-the-blank, and true-false items, the tests are intended to reflect learning of factual information and skills presented in training (Herzog, Trenholm, & Whiteneck, 1974; Nosin, 1975; Quay & Johnson, 1976; Scott, Cienek, & Evans, 1977; West Virginia University, 1970). When the program design includes programmed instruction modules or use of workbook materials, the spot-check tests included in these instruments may double as summative training assessment tools and provide feedback on areas of learning difficulty warranting immediate attention (Smith, et al., 1976; SVO).

Tests used to determine attitude change resulting from learning are often standardized psychological personality or attitude inventories that purport to measure dimensions such as self-other perceptions and self-confidence, authoritarian traits (Herzog, et al., 1974; Nosin, 1975), humanistic orientation (Harris & Harris, 1972), or empathic awareness (Grzegorek & Kagan, 1974). Tests specially designed or adapted to reflect correctional concerns attempt to gauge attitudes about job satisfaction in corrections work (Nosin, 1975; Scott, et al., 1977), the use of punishment (Paddock & McMillin, 1972; Scott, et al., 1977) and coercion (Brutvan, 1969), views of institutions and prison behavior (Paddock & McMillin, 1972; Scott, et al., 1977), and ideas about crime and its causes (Adams, Tabor, Baker, & O'Neil, 1970).

Like opinion questionnaires, such tests can indicate only immediate effects of training. (This is assuming the tests are valid instruments, i.e., the learning objectives of training correspond to those assumed by the test. Correctional evaluation reports often lack information establishing the reliability and validity of the measurement instruments used.) Pre- and posttests focus mostly on attitude change and learning in the classroom, again not necessarily generalizable to the job. As such, effects of social and organizational variables (see Chapters V and VI) and other contingencies heavily affecting job performance, are not recognized. In addition to these limitations associated with pre- and posttesting, it has been suggested that administration of the pretest may influence subsequent learning in training by guiding--or misguiding--the trainees' focus on material presented during the session. For example, a pretested trainee might pay special attention to certain material in the training only because he knows it is covered in test items (Goldstein, 1974).

While traditional correctional training evaluation consists of trainees completing a questionnaire at the close of a session or administration of pre- and posttests to measure classroom learning, there are exceptions. Some studies have gone beyond these standard measures of reaction and learning to determine the impact of training in terms of behavior change and results.

Behavior Observation and Rating

Studies of training programs that attempt to improve human relations, counseling, and behavior modification skills, have focused particularly on assessing training impact in terms of behavior change. Supervisors, offenders, and specially trained observers have used checklists, scales, and performance ratings to report behavior on the job before and after training. Depending on the types and goals of the training, a number of behavioral indices can be used to reflect changes in performance. Raters have kept track of officer-inmate interactions--their frequency, the initiator, tone of the exchange, number of trained skills the officer uses, frequency of reinforcing or positive responses, and so on (Maloney,

Phillips, Fixsen, & Wolf, 1975; Smith et al., 1976; Witherspoon, 1971). When on the job observation was not feasible, judges scored trainee performance using counseling skills with confederate inmates in videotaped practicum exercises (Burkhart, Behles, & Stumphauzer, 1976; Grzegorek & Kagan, 1974; Hosford, et al., 1975). With or without the use of specific scales, trainees' supervisors have used periodic organizational performance reports to assess behavior, have monitored the quality and content of reports written by trainees, and have then communicated performance problems and difficulties to training staff (Nosin, 1975; SVO). Records containing information about legal suits and charges brought against an employee by the clients have been monitored in the assessment of training programs on legal issues in corrections (SVO).

It is crucial that selected behavioral indices and measurement criteria reflect performance changes resulting from training, if an evaluation conclusion is to be valid. Assuming an appropriate measure has been found, it is affected by a wide range of variables in the organizational environment. Use of behavioral observation and rating may offend those who are under scrutiny and resistance may contaminate results. Trainees under observation may behave artificially, simply because they are receiving extra attention. Often observers or raters need to be specially trained to properly use rating instruments and reach agreement on behaviors that warrant notation. If organizational personnel are to be the observers, one must be assured of the authenticity of supervisor and inmate ratings or reports. Conscious and unconscious biases that skew scores either positively or negatively are difficult to detect and filter out of results. The "halo effect" occurs often with supervisor ratings--if the supervisor likes one aspect of the trainee's performance, there is a tendency that the trainee will be rated high on all counts; the reverse may be true also. The same applies to ratings of trainers. In the evaluation methodology addressed later in this report, we try to control such contaminating variables.

Offender Change

While program assessment in terms of changes in staff behavior is most common, the value of personnel training in corrections can be determined also by measures that indicate change in client (offender) behavior. Recidivism is the most popular index of corrections success (Matthews, 1973) (which is unfortunate because of the ease of statistical manipulation and tendencies toward inaccuracy in its use). Records of client or inmate rule infractions, as measured by the number of disciplinary slips or written offense reports submitted over a period of time, are measures used fairly often (Horne & Passmore, 1977; Katrin, 1974). Another behavioral measure includes use of the Jesness Behavior Checklist, an instrument developed specifically for rating the behavior of male juvenile delinquents. The instrument was designed for use by counselors having no training in the use of psychological instruments (Sharp, 1974). On a more abstract level, offender change related to self-perception, anxiety level, and social adjustment have been indicated by scores on standardized personality and attitude inventories (Katrin, 1974; Sharp, 1974).

It is difficult to be definitive in attributing inmate behavior change to staff training programs. Follow-up data on released offenders are especially hard to obtain. Even when extensive data are available, linking effects to training becomes tenuous because of the host of unknown or unmeasured influences external to the training operation in particular and the correctional system in general.

Organizational Change and Climate

Some of the most elaborate evaluation efforts have striven to determine the results of training through measurement of organizational occurrences. The more common approaches are attempts to link training to employee absenteeism, turnover, and tardiness. Records of the frequency of staff-inmate confrontations and physical force required for control have been used as indices of training effectiveness (SVO). Some in-

stitutions monitor critical incident reports that provide information about unusual events, inmate suicides, inmate assaults, confiscated contraband, and escape attempts (Blum, 1976; SVO). Still, others have looked at the relationship between staff members (Brutvan, 1969), particularly line level staff and administrators (Horne & Passmore, 1977), after human relations skills training. The occurrence of procedural changes suggested by ex-trainees or implemented under their direction has been used as a success measure for training programs designed to result in organizational change. When trainees are required to write "action plans" proposing changes they wish to implement upon return to the home organization (see Appendix A, NIC Jail Center site visit report), the extent to which these objectives have been accomplished is used as a rough gauge of training effectiveness (Quay & Johnson, 1976; SVO). The Correctional Institutions Environment Scale (CIES) also has been used to assess staff and inmate perceptions of differences in institutional climate before and after training (Katrin, 1974).

The CIES focuses on three major dimensions in measuring the social environments of both juvenile and adult correctional programs. The Relationship Dimension attempts to identify the type and intensity of personal relationships among institutional residents and between residents and staff. Treatment Program Dimensions assesses the type of treatment orientation the institution has developed (practical preparation for release vs. an emphasis on self-understanding and insight into personal problems). The last scale focuses on System Maintenance Dimensions, related to keeping the correctional unit or institutional functioning in an orderly, clear, organized, and coherent manner (Moos, 1974).

Once again, coming up with legitimate measurement variables--in this case, indications of organizational climate--and linking them to training is difficult. The further removed from training the evaluation focus, the more complex becomes control for intervening forces. As with each succeeding stage there comes an increasing number of variables affecting performance. For example, staff members may tighten up security and attend more carefully to report writing in response to policy and

procedural changes, salary adjustments, change in administration, events such as inmate unrest or a riot threat, or legal challenges in the form of grievances or lawsuits filed for violation of offender rights. If there are such occurrences during the course of training and shortly afterward, resultant employee behavior changes may mistakenly be attributed to training. An awareness of possible contaminants and choice of evaluation design can help account for external variables and make staff more confident of the validity of evaluation findings.

Selecting and Developing Measures of Effectiveness for Performance and Training

Many evaluators feel that assessment of training program success should focus on observable change in trainee behavior, instead of change in attitudes. While reaction and attitude measures are easier to obtain, less is known about attitudes and personality factors, which tend to be more ill-defined and subjective. Behavior is more visible, more certain, and understandable (Mager, 1972). Use of behavioral measures will reduce ambiguity and uncertainty in interpretation of training results and allow for more objective linkage with job competence and performance. Given the variety of measurement variables, the critical point to remember in the conduct of an evaluation is that the relevance of the criteria selected is the fundamental requirement (Goldstein, 1974). In other words, measurement variables must be carefully selected to reflect the job performance the training intends to affect, and these criteria must represent change that is a direct result of training implementation. Since most measures of success suffer from some weaknesses, it has been suggested that the surest evaluation strategy is to use multiple and repeated measures. This has been done frequently in correctional training evaluations (Adams, et al., 1970; Blum, 1976; Brutvan, 1969; Grzegorek & Kagan, 1974; Harris & Harris, 1972; Herzog, et al., 1974; Hosford, et al., 1975; Katrin, 1974; Morton, 1975; Nosin, 1975; Quay, et al., 1976; Scott, et al., 1977; Smith, et al., 1976; West Virginia University, 1970).

Figure X-1 provides a summary of measures that have been used in correctional training program evaluations. For the purpose of organization, the criteria are grouped into four basic stages or steps of the evaluation process (Catanello & Kirkpatrick, 1968; Kirkpatrick, 1977). With this approach, each step assumes an increase in value of the information produced, an increase in the difficulty of assessment, and a decrease in frequency of use as an evaluation strategy (Newstrom, 1978).

PREREQUISITES TO TRAINING (AND EVALUATION)

If, prior to development of the training program, an assessment of job requirements and training needs has been conducted to identify requisite instructional components of the training curriculum, the performance objectives and standards established can provide criteria of effectiveness for training and performance evaluation. Site visit experience and survey results, however, lead us to believe that correctional worker performance levels and needs are rarely addressed during the initial design phase of training. In fact, we suspect that most training evaluations will reveal immediately the absence of such activities. With this in mind, the next portion of the chapter includes a brief discussion of the need assessment and job analysis process. Since primary information about particular techniques and methods is available elsewhere, the intent is to convey only a basic understanding of activities involved.

Assessment of Needs

Before the development and conduct of any training program, the training staff member, along with other organizational representatives must assess training needs. This task involves looking carefully at the work to be performed, the people who are to do the work, and the organizational context within which these activities take place. Needs are identified by comparing acceptable performance with existing performance demonstrated by incumbent employees or staff on the job. Employee job descriptions should include performance requirements and standards that detail acceptable performance and allow for objective assessment of behavior. Observation and assessment of employee behavior and abilities can determine the true or existing level of performance. The discrepancy between "what is

Type and Level of Measure	Method	Measures	Examples
Reaction: Feelings, Satisfactions	• Questionnaires • Interviews • Daily Logs	Respondent (trainee, observer) comments and ratings: • Quality of instruction • Adequacy of facilities • Suitability of training materials • Course structure and content	(Sample Items on Tests, Questionnaires, Rating Forms) Use the following scale to rate each aspect of the program: Excellent — Good — Fair — Poor • Attitude of instructors toward trainees • Usefulness of classroom instruction • Construction and fairness of tests What part of the program did you like best/least? Why?
Learning: Acquisition of new knowledge and subsequent attitude change	• Pre- and Post-tests	Differences in respondent scores on pre- and post-tests and inventories designed to reflect: • Learning of facts and information presented in training	The most common contraband item is usually: a. Money c. Drugs b. Knife/weapon d. Can't say for sure T__ F__ The maximum effective range of a .38 caliber pistol is 50 yards. A helpful response is one that: a. Takes into account the person's record b. Reflects how a person feels about his or her problem c. Always offers a socially acceptable solution d. Can only be made by a professional counselor T__ F__ We can only modify the behavior of people we are acquainted with.
		• Attitude and personality factors--aggression, authoritarian traits, empathy, self-perception; attitudes about punishment, coercion, corrections	T__ F__ No leniency should be shown to convicts. T__ F__ It is wrong for society to make any of its members suffer. Circle one of the two items you feel best describes yourself: A. You want badly to "belong" B. You don't care whether you "belong" or not A. Yours is quick and ready sympathy B. You are stern
Behavior: Change in observed behavior; job performance	• Observation (pre-post) • Rating • Analysis of records and reports	Rater scores and records reflecting behavior frequencies: • Judges' ratings--frequently trainee uses skills taught in training; anxiety behavior during counseling sessions; correctional officer--inmate interactions--initiator, purpose, type, tone • Offender ratings--officer's concern for inmate welfare, fairness; general calibre Personnel records--absenteeism, tardiness, turnover. Supervisor performance reports	Observer record form: (✓) ___ Attentive listening--eye contact, posture, nods head ___ Clarification and feedback--repeats, summaries problem ___ Options and referral--recommends sources for assistance Probation/parole officer performance report: behavior frequency • Deals appropriately with caseload behavior problems 0 1 2 3 4 • Completes paper work neatly and promptly 0 1 2 3 4
Results: Broader, extended changes in clients or system; consequences related to organizational objectives	• Questionnaires • Tests and inventories • Observation and rating • Analysis of organizational records	Organizational climate--agency goal orientation; staff-inmate relationship, staff-administration relationship, morale: • Critical incident reports--contraband confiscated, escape attempts, unusual occurrences • Policy, procedure changes in agency, institution Offender change--recidivism, # rule infractions, disciplinary reports, suicide attempts, self-perceptions, participation in special programs	Which of the following items accurately describe your unit? (✓) ___ Activities for residents always follow the same daily routine ___ Residents have a lot of say in what goes on ___ The facility is always neat and organized ___ Staff frequently argue among themselves

Figure X-1. Training Evaluation--Types and Levels of Measurement

wanted" and "what is" indicates areas of need (Anderson, Ball, Murphy, & Associates, 1975).

First Steps

The next step of the assessment involves making a decision as to whether the need is an issue that should be dealt with through training or by way of some other intervention. Training cannot be a cure-all for all organizational and personnel problems. Some needed performance changes may come about only through implementation of new employee selection and hiring practices, or policy, personnel, or operational changes within the organization, or restructuring of the organization. A good question to keep in the forefront: Is the organization ready for the changed man or woman? For changes in behavior on outlook brought about by training to take hold a change in the organization's way of doing things is often called for. If the organization can not answer, "yes" to the above, it had best think again about the desirability of offering the training in question.

The methods of obtaining information for a need assessment are the same as those utilized in program evaluation. Table X-1 provides a summary of available information sources, methods of study, and some of the weaknesses of each approach. For additional explanation of the measures and their limitations, refer back to the discussion of measures used in an evaluation program.

Finally, the assessment process includes prioritization of needs as elements that are to be in the training curriculum. There are several guidelines that the trainer can rely upon in deciding the importance of needs:

1. Learning difficulty of the task:
   - Is the behavior required by the task hard to learn?
   - Can the skill or knowledge be picked up on the job?
   - Does it take an instructor's guidance and supervised practice to acquire the skill or knowledge needed?
   - Can task activities be communicated through training?

Table X-1

Obtaining Information for Job Analysis and Assessment of Training Needs: Measures and Methods

INFORMATION SOURCES AND MEASUREMENT PROCEDURES	RESERVATIONS--LIMITATIONS OF EACH METHOD
Interview--with employee, job incumbent, supervisor	• To what extent are findings influenced by personal interaction and communication between parties?
Questionnaire--open-ended, structured, checklist (completed by employee, supervisor)	• Quality and utility of instrument • Validity of responses
Meeting with Technical Personnel to identify goals, standards, needs--organizational managers, administrators, personnel representatives	• Are upper echelon personnel familiar with on-line events and problems?
On-the-Job Behavior Observation--by analyst, supervisor	• Observer interference with job activities • Observer bias, contamination
On-the-Job Participation; job analyst performs the work	• Interference with work activities • Ability of analyst to do the job • Is a true sampling of total work experience obtained through participation?
Diary Method--employee maintains daily log of work activities	• Interference with work activities (extremely time consuming) • Willingness and ability of employees to carefully record actions
Critical Incident--observation or inquiry focuses on instances of behavior that are characteristic of satisfactory, or unsatisfactory, employee performance	• Are essential job behaviors overlooked?
Psychological Measures--tests, inventories, rating scales	• Can job-relevant tests be obtained or created? • Complications (validity, adverse impact) associated with psychological testing
Review of Other Documentation Sources--organizational policy and procedure, training manuals, personnel records, prior research and technical literature, job descriptions and standards prepared by commissions and professional associations	• Quality, generalizability and dating of studies, reports, and available information

2. Frequency of task performance/time spent performing the task:
   - How often does the worker need to use this skill/knowledge?
   - How much of the worker's time is spent on this aspect of the job?
3. Number of people (workers, incumbents) that demonstrate the need:
   - How many individuals require training for this performance?
   - Is the training required the same for all?
4. Criticality/importance of the task:
   - Does having or not having this ability distinguish between a good or bad worker?
   - What would happen if the worker could not perform the activity competently?
     - personal injury risk involved
     - costs incurred by reassignment or reallocation of other workers
     - possibility of damage to property and equipment
5. Organizational readiness for new or changed task performance:
   - Does the organization support this activity?
   - Are personnel willing to accept this responsibility?

Pre-Training Analyses of Needs

Correctional training practitioners have, to a limited extent, conducted preliminary analysis and assessments of training needs. These efforts usually have been informal and intuitive. The more formal methods employed by training staff members to identify details of the job and training needs are represented in Table X-1. As one would expect in light of evaluation endeavors, the most commonly used tactic of determining needs is to ask the employee, "In which areas do you think you need training?" or to circulate within the agency a checklist providing

options representing possibilities for training. The return rate of such questionnaires, however, is low and discouraging for the trainer who attempts the assessment (SVO).

Unit supervisors and managers may be asked to identify performance difficulties and training needs of their subordinates. This information is also gathered by means of questionnaires/checklists, or employee performance reports are analyzed to identify areas in which performance might be improved through training. Finally, committees and advisory committees made up of department heads, administrators, and managerial staff are established to identify training needs.

Trainers may visit the job setting or visit facilities to observe work. The trainer-observer should adopt a systems perspective in viewing activities in order to pinpoint vulnerable areas and bottlenecks associated with work difficulties or particular organizational concerns. Trainers may also go out onto a job periodically and participate in work activities. This technique is especially feasible and appropriate in corrections, since many trainers have risen through the ranks and are capable of performing the work adequately. The practice ensures that the trainer is in touch with updated organizational policies, procedures, and new practice techniques. Knowledge of job components and special needs may also reach the program through a kind of osmosis. A trainer who has been transferred into the position from a line level job should be able to rely somewhat on his own job experience and perceptions in selecting instructional components.

Trainers most frequently solicit and rely heavily upon feedback obtained through informal networks and contacts, and the information gathering process is not formally recognized as a need assessment per se. A supervisor may casually call up a trainer to report that the last class of trainees is having difficulty correctly completing parole recommendation forms; former trainees may drop in on a class to enable the trainer to know how things are going--that the way in which they have been taught to conduct counts is not how counts actually are done in the prison. This

informal exchange between training staff and practitioners in the field yields information that has most pervasive and influential impact upon training and determination of needs.

While these informal assessments of training needs are useful to training personnel, they fail to provide complete and specific information about performance gaps. It was noted earlier, that the most effective manner in which to identify training needs is to compare true employee behavior with performance requirements and standards laid forth in job descriptions. One reason why this approach is seldom used is because few correctional organizations have developed adequate job descriptions that include standards for assessment of personnel behavior. The final portion of this chapter contains a discussion of job analysis as a tool to serve this purpose.

Job Analysis

There are no universal prescriptions that guide the conduct of job analysis. Here too, the process varies according to the purpose and needs of the program, the resources and time available, and the type of technique applied. Job analysis refers to any number of activities carried out essentially to define what a job is and what employee behaviors are demanded by the job. The analysis may be undertaken for several reasons--to help establish valid employee selection procedures, to aid in the development of performance standards for periodic employee performance evaluations, or to determine content needs of staff training programs.

One can conceptualize job analysis as a systematic study that entails investigation of three general components: (1) the organization and its environment; (2) the work itself, comprised of a number of activities and work task clusters that may be referred to as jobs; and (3) personnel, being the workers or prospective employees of the organization. The analysis is aimed at understanding how these three dimensions of the total system function and interact to influence job requirements and training needs.

The Organization

Relevant characteristics of the organization include agency goals, resources, and internal and external factors of the organization and work environment. Goldstein (1974) addresses in detail each of these components as it influences training outcome and subsequent employee performance.

Organizational goals and training focus should agree, if the trainee is to perform acceptably on the job. It is of particular importance that correctional training staff attend to organizational philosophy and goals, since corrections, as a field, has demonstrated little enduring dedication to a common mission or consensus on the purpose of services provided. Training participants interviewed during site visits frequently expressed a desire that their bosses be required to attend training sessions. Goldstein offers this very statement as an indicator that training program values differ from the approach of administrators and upper level supervisors.

If the goal of an institution is solely to incapacitate the offender or to uphold retributive justice, a staff training program that emphasizes topics such as offender rehabilitation, treatment services, and the development of counseling skills invites serious transfer problems and sets the trainee up for a confusing and frustrating experience on the job. Successful transfer of divergent training segments can occur, however, if the content is integrated carefully into the broader organizational plan and skills taught are presented as means for achieving the organizational goals. Crisis Intervention training, for example, endorses non-violent, empathic, helping behaviors for staff, yet sidesteps controversial rehabilitative ideals by basing its approach on the goal of providing a safer, more controlled work environment for prison staff members.

Organizational policy and training content must reflect current policies and procedural directives, and (hopefully) be related to organizational goals and objectives. One institution surveyed in this study indicated that a course in their training, "Press Relations for

Correctional Officers," was a "less useful" course as compared with other courses. The explanation for such an assertion was that institutional policy prohibited correctional officers from providing any information to the press. Another respondent identified "auto traffic control" as a "less useful" training component--training provided to correctional officers who had no job responsibilities requiring direction of traffic. Job analysis and need assessment, including review of organizational policies, should prevent the development and inclusion of unneeded training modules.

Assessment of organizational resources entails gathering information on manpower capabilities, financial support, and characteristics of physical facilities. Questions that address relevant manpower issues and cost factors would be: Are employees demonstrating any specific performance difficulties? Would it be more cost-effective to hire a new staff member than to continue paying excessive overtime costs? How many replacement staff are needed to cover for personnel attending training? Records of absenteeism, disciplinary actions, and tardiness may reveal facts that warrant attention through training or other intervention.

Appraisal of physical resources includes a complete inspection of facility design and maintenance characteristics to identify information that should be included in training content. Training can be used to alert a new employee to design flaws of an institution, e.g., areas in a jail not adequately constructed to prevent inmate suicide attempts. Similarly, blind spots or escape problem areas particular to the layout of a facility can be identified for a new correctional officer preparing for tower duty. The job analyst also conducts an inventory and determines the condition of tools, equipment, and work aids used on the job. Here, especially, one objective is to keep training relevant and up to date. Agency use of computers for budgeting purposes, for caseload classification in probation and parole, for storage and retrieval of community resource and referral information is rapidly on the increase.

Organization plans for the purchase of and conversion to computer systems can be accompanied by foresightful recognition and analysis of new training needs. More frequently, a change in report forms and paper work requirements will necessitate modification in training content.

Finally, an awareness of human, financial, and physical resources within the agency must be accompanied by knowledge of support services provided by the surrounding community. In a community where residents and business operators are agreeable to the implementation of youth work release and employment programs, innovative staff training may include educating trainees about their role in relation to such services (i.e., training in public relations skills, community resource use, or pre-release preparation of offenders).

Internal factors refer to psychological, social, and physical conditions of the work place. They include less formal characteristics of the organization such as leadership provided personnel, nature of supervision received, staff morale and teamwork, reward and motivational aspects of the unit, and quality of staff-administrator relationships. Informal peer group norms and their influence on work performance are addressed in this category of analysis. Such investigation can lead to recognition of special features of the job context such as the extent of job pressures and stress placed on the employee. This area of examination is especially important in correctional institutions, where individuals live and work in a closed, volatile environment. An extensive analysis might include efforts to identify niches or specialized subenvironments within the prison, that are separate and distinct climates, their use related to particular staff members, inmate groups, and specific activities (Toch, 1977). Such spaces that offer varying degrees of quiet, safety, support, or activity might be presented to the correctional officer as informal resources available to help reduce stress and dissatisfactions within the prison.

Physical conditions under which the work is performed are also described with attention to factors such as work location; temperature variations, illumination, proximity of workers and clients, noise levels;

and other characteristics of the setting, such as hazards particular to the job. Any ecological characteristics of the work place that may potentially alter job behavior and individual activities should be recognized in the analysis (Dunnette, 1966).

External factors encompass broader dimensions such as the pervasive social, economic, and political conditions that affect the job, the worker, organizational activities, and decision making. While such variables are often powerful, yet not obvious (and may, in fact, remain beyond control of the training staff member), they must be recognized and dealt with head-on in program development and evaluation efforts. Part C of the report addresses these external factors in detail.

The Work

The second major focus of the job analysis process is the work itself. For purposes of study, the term "job" refers to a cluster of similar work activities carried out to achieve some relatively stable common purpose (an organizational goal or objective). The job analysis, however, is directed at the level of work task and the fundamental job component itself, in order to assure attention to specific behaviors and performance requirements (Dunnette, 1966). Work components are also referred as positions, activities, elements, work behaviors, and the like, depending upon the level of work being addressed. Some breakdowns illustrating the possible focus of a study follow:

Work----------> Job----------> Task
Job----> Position----> Task----> Element
Occupation----> Job----> Position----> Task

The degree to which each task is dissected or broken down into task components or elements varies according to the preference of the analyst (Sparks, 1979). For example, note the following task description derived from a job analysis of tasks included in correctional worker counseling activities (Fine, personal communication; National Advisory Network, 1976).

While the description is presented as a single work task, another analyst might separate each different behavior and claim that five distinct tasks are involved.

> TASK: [The Worker] *advises/counsels* family/wife of offender up for parole in relation to reentry problems, *listening, asking questions, reflecting feelings, and suggesting ways of coping* with problems and anxieties, (according to own knowledge and experience) in order to help adjust to problems of both family and parolee.

This part of the job analysis concerns itself with what is done and how the work is performed. It involves listing tools, equipment, or work aids used on the job, and providing information on techniques and procedures employed in carrying out work activities. The object of worker functions may be specified--different job analysis techniques offer various ways of categorizing and recording job relevant information (McCormick, 1976). Under the Functional Job Analysis (FJA) approach, activities involve dealing with data, people, or things (U.S. Department of Labor, 1972; Fine, 1971). Data tasks concern the processing of information; data oriented work for example, might involve those tasks performed by a staff member in a reception and diagnostic unit, compiling background information and computing test scores of incoming offenders. These functions involve some type of cognitive action such as comparing, analyzing, writing, and coordinating. Many functions of a unit supervisor or counselor involve dealings with people, including tasks that entail use of interpersonal skills such as listening and responding, giving instructions, help, or services. A task such as the operation of mechanical gates from a central control room or an activity involving the use of weapons would be classified as a function involving things. Such functions refer only to the physical handling or manipulation of objects and equipment.

The result of a job-task analysis is a task description, which provides a brief, explicit statement of what the work entails. While there may be some variation in content, a typical task description delineates worker actions required by the job, the object of task activities, and

expected results. For purposes of training and evaluation, the task description may be supplemented by a specification of performance standards and corollary training content needs. The performance standards--criteria by which performance will be measured--must be presented in quantitative and/or qualitative terms. The former allows for objective measurement of behavior by means of numerical standards that need no interpretation. The latter are less specific and provide for assessment in terms of more subjective, yet informative criteria. The training content may also be classified in two ways. First, as needs of a more general, functional, or basic nature; second, as specific task related requirements likely to be learned through on-the-job experience or advanced training (Fine, personal communication). See Figure X-2 for an example of a task description.

Personnel

The final category of analysis involves a description of personnel requirements in terms of skills, knowledge, and personal attributes demanded by the job. Here, there is an effort to match human traits such as learning ability, aptitudes, temperaments, interests, and physical capabilities with performance requirements established in the task description. A job analysis that contains an in-depth coverage of this component probably provides for recognition and classification of individual differences along the dimensions listed below:

- Aptitudes represent individual tendencies and abilities that may influence learning capacity. This component can include a number of factors (measurable through psychological testing), such as general intelligence, numerical ability, verbal or mechanical reasoning capacity, and physical capabilities related to coordination, perception, and dexterity. Distinctions among different aptitudes vary--numerous multiple aptitude batteries have been developed for use in educational testing, counseling, and personnel classification (Anastasi, 1976), and each test varies somewhat from the others in which aptitude factors it identifies.

- Some analysis procedures specify training time as distinct from aptitude factors (U.S. Department of Labor, 1972). This component refers to the amount of general educational development (GED) or specific vocational

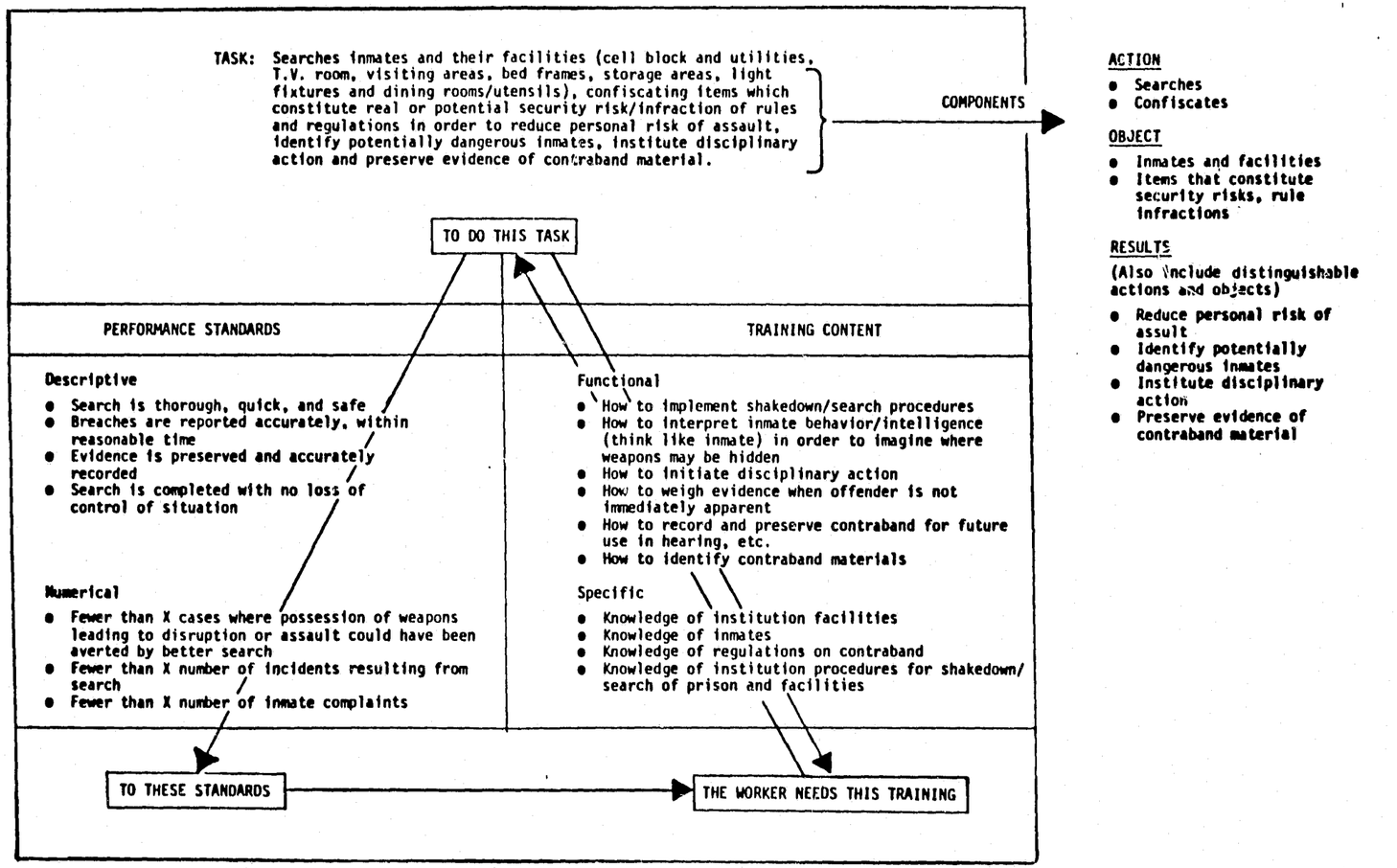

Figure X-2. Example of an FJA Task Description
(modified from National Corrections Task Bank, 1976)

preparation (SVP) an individual will require to adequately perform the job in question. The former deals specifically with the trainee's reasoning development, ability to follow directions, and capacity for acquiring basic language and mathematical skills. On the other hand, SVP is one estimate of how long it will take the trainee to learn techniques and absorb information needed in the specific job situation. There is here, an inherent assumption that prior life experiences--education, hobbies, work, or training--may influence individual performance along these lines. Such information is often requested on job and training application forms.

- "Personal" traits or temperament are indicators of how an individual may adapt to various job situations. Temperament (characteristically, the innate aspects of personality) may be defined for a job analysis in an attempt to identify or predict how one will react to different demands of the job that call for dealing with people, accepting responsibility, using discretion, or performing under stressful circumstances.

- Worker interests, like temperament, are closely linked to personality theory, and deal with emotional needs and highly subjective personal attributes. Individual differences in interests relate to whether one may prefer to work with people or things, can perform better in a highly structured or less structured environment, or if one has need for a job that will lead to personal power, influence and prestige.

- Physical abilities demanded by the job are, of course, more observable and objectively measurable items. This component of the job analysis specifies requirements of the work such as physical strength and flexibility, balancing and reaching abilities, and need for sensory faculties such as sight, hearing, smell, and touch.

This area of analysis touches on the sensitive issues of measurement of personal traits that is subject to regulations delineated in Equal Employment Opportunity Commission Guidelines. These legal restrictions, however, are related particularly to selection, promotion, and placement decisions made on the basis of test scores. Of particular relevance to training representatives are the learning-related measures that may affect decisions regarding training processes (at what speed can individuals in the class learn, how should instructional modules be spaced) and training

content (at what level of complexity and detail should information be provided, are trainees already more adept in one area than another, i.e., weapons use vs. communications skills).

Psychological characteristics addressed in job analysis may be rated, scaled, and assessed through interview procedures, or more often, through testing. Measurement of less observable, less tangible human qualities like aptitudes, interests, and temperament is a highly technical and complex undertaking; difficulties associated with even the most well-established psychological batteries and inventories render their use subject to debate. Even when measurement tools are accepted, procedures for determining validity are lengthy and time consuming. Validity concerns and problems involved in collecting information for all areas of a job analysis are generally the same as those presented earlier in the discussion of evaluation measures (see Table X-1).

Job analysis and need assessment have been singled out in this chapter as necessary components in the development and design phase of a training program because our findings have indicated this area to be a real trouble spot in correctional training. Trainers invariably identify as critical factors impacting upon training program effectiveness, (1) whether or not training content is relevant/appropriate in light of job responsibilities, and (2) whether the level of instruction is congruent with trainee knowledge and abilities. If efforts to answer these questions through analyses have been made during the preliminary stage of training, an evaluator can return to the initial performance standards developed and use them as criteria for training effectiveness (assuming that training effectiveness is related to the quality of employee performance). If job components and training needs have not been identified prior to program evaluation, job analysis and need assessment then become tasks to be conducted as part of the evaluation process.

OVERVIEW

We have discussed motivations for conducting an evaluation, difficulties involved in initiating and carrying out program assessments, strategic components of the process, and some commonalities of correc-

tional training evaluations. The problem issues, as we have conceptualized them, are summarized below.

Reasons for resistance to evaluation:

- Professional evaluators lack credibility and familiarity with in-house circumstances.
- Failure of evaluation information to have significant impact upon decision making.
- Evaluation results threaten staff jobs and program funding.
- Confusion in selecting and implementing the appropriate evaluation strategy.
- Lack of time and funding to plan and carry out evaluations.

Characteristics of evaluation efforts:

- Conduct of evaluations primarily by consultants external to the organization (high cost and concern with lack of organizational insight).
- Primarily ex post facto assessments--little control over non-training variables influencing performance.
- Tendency to attend to wants instead of needs of trainees.
- General failure to assess transfer of learning and impact of training on job performance.

In light of increased legal intervention and regulations, emphasis on new accreditation requirements, and the growing sophistication and professionalism in the field of correctional training, there is an increasing need for competent and complete evaluation of correctional personnel training programs. For many, this process may involve starting at the program development stage and conducting job analysis and need assessment.

The comprehensive evaluation model described in the next chapter has been designed to provide a basic structure and guidance for selecting and applying training evaluation methods and procedures. Many of the issues we have discussed so far are related to the use of evaluation and to the attitudes with which people approach the evaluation process, and are thus tangential to methodological and technical concerns.

We have mentioned that evaluation is a process heavily influenced by values, politics, and personal beliefs--variables often difficult to assess and control. So often, members of training staffs do not have the influence to achieve assistance and cooperation from the larger system. We have sought to identify and discuss some of the issues surrounding evaluation so that they may be realized in evaluation planning, in order to increase the possibility that evaluation efforts will be accepted and supported. Recommendations and reminders of strategies that the staff member conducting an evaluation can use to facilitate the process accompany the presentation of the model in the following chapter.

CHAPTER XI

AN EVALUATION STRATEGY FOR CORRECTIONS

INTRODUCTION

In the last chapter we discussed a number of topics relevant to planning and conducting an evaluation of training. In this chapter, we will describe a specific strategy for carrying out such an evaluation. This strategy is not meant to be particularly innovative, or technically sophisticated, though it is a distillation of much of the current thought in evaluation research. It differs from other evaluation models more in purpose than in content. It is a strategy that trainers, administrators, and personnel representatives, not totally familiar with evaluation methodology, research design, or statistics, can use to evaluate their training program.

This evaluation process is outlined in Figure XI-1. It is composed of six phases. We will discuss the decisions, concerns, and actions needed at each stage.

IS THERE TO BE AN EVALUATION?

The first question to be asked in planning an evaluation is not whether a training program is to be evaluated, but whether a planned, systematic evaluation is to be attempted. Every program is subject to informal judgments, but the informal evaluations provide no control or measurement of the completeness, bias, or validity of the evaluation. The first issue then centers on whether an intentional evaluation, with rules, procedures, and standards will be conducted.

Although evaluation is generally advisable, it is not always the best course of action to carry out a formal evaluation. Three situations can be specified where it would be inadvisable to perform a formal evaluation; these are instances in which the evaluation is either cost-ineffective, irrelevant, or impossible to perform (Anderson & Ball, 1978). An evaluation effort may prove cost-ineffective if the training program is

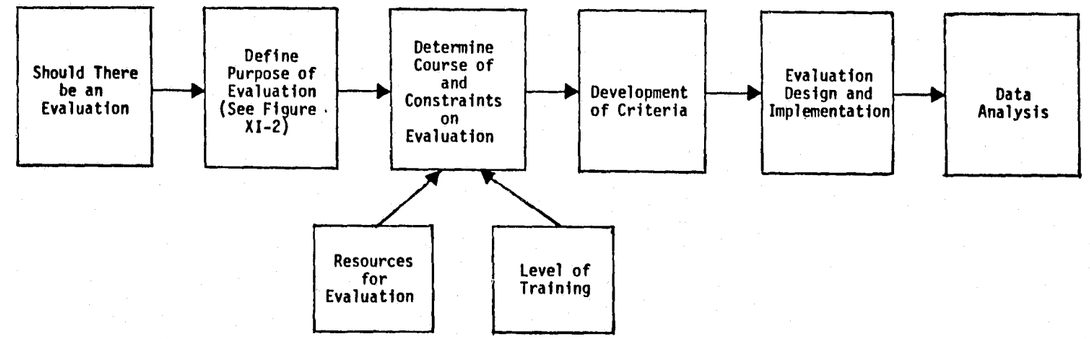

Figure XI-1. Steps in the Evaluation

a one-time effort, if it is provided to very few people, or if the evaluation costs more than the training. An irrelevant evaluation will result, if the information derived from the evaluation does not influence decisions about training. Finally, professionally responsible evaluations cannot be performed in settings with no financial resources or competent personnel to perform the evaluation.

DEFINE PURPOSE OF EVALUATION

Once it has been decided that the training program is capable of supporting an evaluation, the purpose of the evaluation must be determined. One may draw an analogy between an evaluation without specified purpose and a training program without specified goals. Neither are likely to produce useful results. It is important to clarify the purpose of the evaluation, since evaluations with different objectives require different emphases and directions, and are limited by different constraints.

In making this decision, two issues have to be considered. First, the evaluation should be concerned with getting information that will be used; that is, information that will have an impact on the decision-making process and have an effect on the training program. For example, deciding to compare the relative cost of two training programs is not appropriate if it has already been determined that one of these will be selected regardless of cost. Evaluating trainee selection procedures is also of no benefit, if the state has legislated mandatory training.

Second, an evaluation should be focused so that it produces data that are useful--data needed to make the most pressing decisions. One has to decide which of the areas that are open to change are most important and/or present the most intense need. For example, should evaluations of the overall effectiveness of the program be undertaken if the program is still being revised?

There is no pat procedure for specifying the purpose of an evaluation. In doing so, one should consider gathering information about prior training, legal requirements, and possible impacts by and upon others

involved in the correctional system. From the beginning of the evaluation, lines of communication should be open to all relevant parties--administration, program directors, trainers, and trainees. Input from these sources will help the evaluator identify those areas of greatest need that are susceptable to being changed.

Evaluation Purposes

The following sections describe three general purposes of evaluation: (1) preliminary installation of a training program; (2) formative evaluation of a program; and (3) summative evaluation of the results of the program. This overview is meant to help the user to identify potential areas of investigation in a coherent manner; it is not meant to restrict the evaluation to an a priori structure or to imply that an evaluation can have only one purpose. The schema presented here benefits from the comprehensive presentations of Anderson and Ball (1978) and Scriven (1974). Topics of concern within the three major evaluation purposes are shown in Figure XI-2. Each will be discussed briefly.

Program Installation Evaluation

There are many aspects of a training program that can and should be examined critically before the program is started. This type of evaluation, frequently referred to as a front-end analysis, encompasses activities that have previously been associated with needs assessment. However, such evaluation not only determines the needs to be met by the program, but also assesses the likelihood that the program will meet those needs.

Given that the organization has decided to develop a training course, it is necessary to identify and prioritize training objectives. An installation evaluation should ascertain whether the training program will, in fact, be addressing the most intense needs of the organization. A needs assessment conducted prior to the installation of a training program, provides information to determine the value of training objectives (see Chapter X).

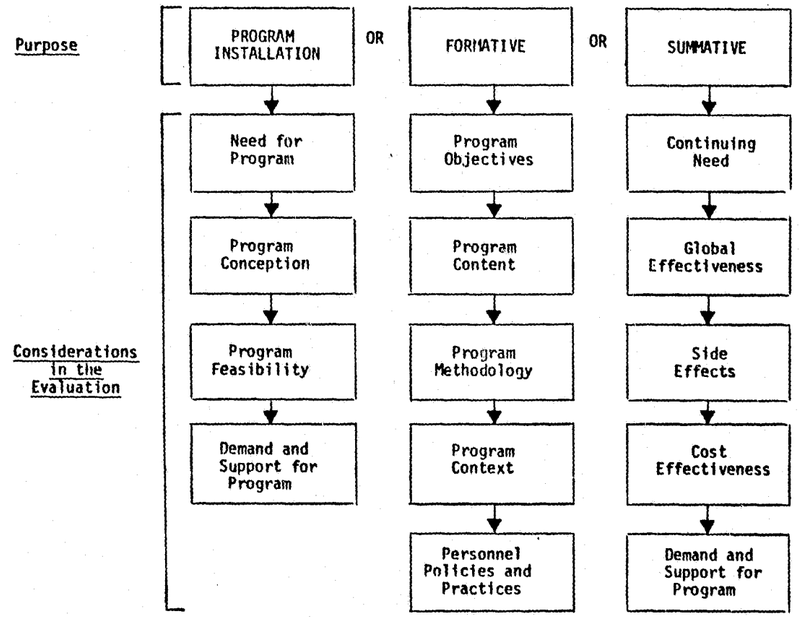

Figure XI-2. Definition of Purpose of Evaluation

An installation evaluation addresses the origin and development of the training program. It asks, how closely related is the content of the program to the needs which must be satisfied? All too often, a specific need is identified and then responded to with canned training programs that are not relevant to the need. For example, a high rate of inmate-officer conflicts may lead to the decision that there is a need for "interpersonal training," which is then met through any program with that title, regardless of whether the training deals with inmate-officer conflicts.

Aside from the important issue of usefulness, installation program evaluations must determine whether or not the training program is feasible. Given even a superbly designed program, one still must ask whether the capabilities--financial, personnel, material, operational--are available to produce the program, and if they can be maintained throughout the life of the program. This includes an attempt to estimate costs of the program and to compare these costs to alternative training strategies.

Along with assessment of the need for training and the appropriateness of the training course for meeting this need, is the assessment of support from the surrounding institutional environment for the behavior that will be trained, or for the training itself. Too often employee training programs are installed without identifying organizational barriers that might make it impossible for trainees to apply their newly acquired knowledge, or without recognizing attitudes and policies of the administration that may negate training effects. Resistance from older officers to on-the-job use of new techniques taught in training is a classic, yet continuing, problem of this type that surfaced during our site visits. The confusion this caused for new employees was compounded when organizational policies failed to reinforce training. Reluctance of supervisors to release personnel for training, or unwillingness by employees to leave their work to attend, are additional training-organizational difficulties that must be addressed and resolved prior to implementation of a program.

Formative Evaluation

The formative type of evaluation is tied most specifically to the Instructional System Operations Model (see Figure III-2), which can be used for this purpose. The main objective of such an evaluation is to provide feedback to improve the training program; this is done by inspecting each of the separate components that comprise the training process. The goal is to specify aspects of the program which relate to its effectiveness at different stages, and suggest modifications. Emphasis is on describing program processes, not program products.

A formative evaluation should cover a broad range of issues. One major concern is the evaluation of the program objectives, which may be examined from two perspectives. First, are the course objectives matched to trainee needs? Second, are these course objectives accepted by those involved? The lack of trainer, trainee, or administrative acceptance of the objectives can reduce the probability of the course reaching its objectives.

Formative evaluations also examine the course content. The content must be relevant to course objectives, and its level must be geared to the trainee population. The level of the course, including the sophistication of the language used and the reading level of training materials, must be appropriate to those participating in training. In addition, one must consider the amount of knowledge the trainees have about the information presented, in order to avoid boring, or overwhelming, the trainees.

Presentation of material is closely related to program methodology--issues apart from course content, such as the length and pacing of sessions, the degree of trainee autonomy, different instructional techniques, the reinforcement system for training, etc. The methodology must be appropriate for teaching the subject matter, as well as interesting for trainees. Teaching skills of the trainers must also be considered in the development of course content and selection of methodology. Trainer's skills (technical knowledge and style) can be major determinants in the success of a course.

Knowledge of personnel policies and practices that effect the trainee population is necessary in a formative evaluation. Who is the training reaching? Are these the people that need and can best utilize the training? Do the trainees have the ability to benefit from the training? With respect to the training staff: Are they qualified to teach the course? Have the trainers received sufficient instruction and guidance to deliver an effective program? Is the system designed to retain good staff and weed out the rest? The effects of personnel policies can not be overemphasized; in many cases it is the training staff that make or break the program.

Summative Evaluation

Summative evaluation represents evaluation in the commonly used sense of the word. It involves determining the effectiveness of a program following its completion. It differs from the other types of evaluations in that it relies somewhat less on judgment and somewhat more on data, e.g., test scores, employee job behavior, offender records, to decide if the program is working. However, the differences are a matter of degree, since summative evaluations must be concerned with some of the issues raised in formative evaluation--the usefulness and importance of the program.

The more usual component of summative evaluation is the assessment of the effectiveness of the program as a complete project. Training effectiveness can be evaluated in a number of ways. Hamblin (1974) specifies five effects of training that could be considered: reactions to the training; learning; changes in job behavior; changes in the organizational processes and functioning; and increases in the achievement of ultimate goals of the organization, e.g., reduced recidivism.

Perhaps the lowest level of an evaluation to be considered is that of reactions to the training program. How does the trainee like the training? Every training program will stimulate some type of trainee reaction (even if it be indifference), and the majority of correctional training program evaluations have included an attempt to measure these

reactions. Unfortunately, most evaluations have provided only such consumer attitudes, without offering any substantial evidence as to the nature of impact, if any, upon performance in operational settings.

The second level at which training programs may be evaluated deals with what the trainees learn. The fact that trainees are expected to learn (defined as acquiring the capability to behave in new ways; not necessarily demonstrating new accomplishments) is what separates training programs from entertainment. In assessing trainee learning one may ask: Can the trainee accurately describe the proper procedure for conducting a count? Is the trainee aware of organizational policies and regulations pertaining to the job? Does the trainee know how to properly fill out and file a disciplinary report? Such measures, once again, reflect learning of training material, not performance on the job.

The third level at which a training program can be evaluated is with attention to the effects it has on job performance. The issue is often called transfer of training. In order to evaluate programs at this level, it is necessary to specify the goals of training in terms of behavioral accomplishments and to collect information about job behavior. Evaluations designed to obtain such information may ask: Does the employee conduct inmate counts correctly (as taught) in performing his/her job? Has the employee violated any organizational policies or regulations in the course of his/her employment? Have disciplinary reports submitted by the employee been completed according to specified procedures?

If the training program does not produce the desired effects on job behavior, the evaluator needs to determine why. It is possible that the training program is not relevant to job requirements, that information was presented poorly and not learned, or that a failure occurred elsewhere--the employees never attended training, poor attendance hampered learning, or events in the work place precluded the training from using the newly acquired skills.

The final two levels of Hamblin's model deal with the organization and the ultimate goals of the organization. Evaluations at the level of the organization ask whether changes in employee's job behaviors due to training affect the functioning of the agency (hopefully, for the better). For example, assume that a training program has been shown to increase personal interactions between correctional officers and inmates--one of the training objectives. How do increased interactions affect the organization? Officers may provide more referral services; violent exchanges may decrease; grievance reports against officers may decrease; or improvements in work atmosphere may decrease absenteeism. At the ultimate level of evaluation, one asks whether the changes in the agency's functioning helped to achieve its goals. These goals usually concern the impact on offender behavior. For example, did the program contribute to successful employment of paroled offenders? Have fewer probation revocations occurred since the installation of training? Have there been fewer inmate escapes since officers take security training?

At these last two levels of evaluation, the number of influences on the criterion or goal are numerous. Training will be only one factor, often one of the less important ones, that influences offender behavior, or other organizational goals such as social reintegration or restraint of offenders. One must decide if it is worth the resources to use evaluation measures that are far removed from the training itself, or if it is possible to change organizational structure and functioning to enhance the benefits inherent in the training provided.

Hamblin recommends evaluating program effectiveness at each level. Such an approach may be prohibitively expensive, require more staff than is available, or be considered not worth the effort. Regardless of how many levels the evaluation addresses, decisions have to be made about how to measure effectiveness at each level.

Effectiveness can be measured within two general frameworks. The training program can be evaluated on the basis of performance standards related a priori to the objectives of the program. In this approach, the consequences of training, whether it is some behavioral change or

change in organizational functioning, is evaluated in comparison to a stated goal. Cost standards and standards of employee performance are two examples of the standards that may be used for evaluation. Often, clear, specific standards are not available. Under these conditions, effectiveness is evaluated by comparison of the performance of trainees after training and performance before training, or a comparison of the performance of trainees after training and the performance of another group, either untrained or trained in an alternative program.

When comparing performance of trainees after training, either to some standard or to the performance of other individuals, one must specify the specific performance that is to be compared. Generally, the performance selected is in line with the training objectives. An alternative view, which opposes narrowing the focus of the evaluation, is that represented by goal-free evaluation (Scriven, 1972). Scriven's concern is that evaluators should take into account all of the effects of a program, and not limit evaluation to predefined objectives only. Scriven views consideration and evaluation of training goals and objectives as an unnecessary and possibly contaminating step, since it may give unwarranted attention to intended outcomes while slighting important "side-effects." In Scriven's view, gathering data on a wide range of actual effects and comparing them to demonstrated needs protects the evaluation from being too narrowly focused.

We take the approach that it is important to focus the evaluation on program objectives for a number of reasons. First, the goals of the program presumably have strong influence on the direction and operation of the training program. Examining the objectives is a necessary part of understanding the dynamics of training. Consideration of objectives provides structure to the evaluation. In a realistic setting, the number of possible effects to be measured is enormous, with the majority inconsequential or irrelevant. Decisions must be made prior to data collection (as a necessary step in allocating the evaluation resources) as to which sources of information will be the best bet. It is not feasible or efficient (and will probably later prove uninterpretable) to

measure training effects using the "fish net" approach, catching anything in sight. In addition, such an approach may prove cost-ineffective, since a great deal of useless information will be gathered. The "Bell-ringer," (1979) a periodic review of the criminal justice evaluation field, recently pointed out that there are clear economic benefits to limiting unimportant, nonessential evaluation activities such as collecting data with no plans for their use, answering questions no one is asking, and answering questions that have already been answered.

Focusing on program objectives will reduce the influence of personal bias. Without reference to program objectives, the evaluator is often left with personal discretion in selecting the criteria and method of evaluation. Such personal biases are rarely articulated, not as subject to criticism by others involved, nor generally as appropriate as are program objectives that are understood and agreed upon by those responsible for carrying them out.

Although we believe that training evaluation should focus on training objectives, this does not mean that objectives only are to be studied. One must also be aware of side-effects of training, and the present model places special effort on identifying both positive and negative unplanned effects of the training program. For example, skills gained in a program aimed at improving correctional officer offender interactions may be used by officers to improve their communication with supervisors. Training, in general, may improve job satisfaction if workers feel more competent and in control of potentially tense situations. Side effects, however, are not necessarily positive. Training may produce tension between new staff and more experienced staff, who are threatened by the new training.

CONSTRAINTS ON THE EVALUATION PROCESS

Once the purpose and direction of the evaluation are determined, one must address the life constraints of the evaluative situation. The environment of the training program has a strong influence on the nature of the evaluation, and plans for evaluation activities should take into account these environmental constraints.

Three major influences in any evaluation effort are the financial, political, and technical support for the evaluation. The available funding will affect many aspects of the evaluation, such as its length and depth, staffing, and methodology. Not only must the amount of funding be considered, but the stability of the funds as well. A short-term evaluation may be more beneficial than a long extensive evaluation that is terminated prior to completion.

Site visit and survey findings revealed that agency superintendents and training staffs frequently have little knowledge of budget allocations or details of expenditures for staff training activities. The undertaking of an evaluation requires, prior to implementation, assessment and planning of program costs. This calls for knowledge of funding particulars (source, amount, apportionment, stability), a detailed written proposal providing for coordination of needs, resources, expenditures, and control of evaluation activities within the available budget. Correctional agencies traditionally do not provide substantial amounts of funding for evaluation--management of available resources is essential.

Political support--the approval of agency and departmental administrators--for an evaluation is essential. Often the level of financial support merely reflects the political state of affairs. The evaluation of training programs requires a cooperative network that not only encompasses the administration and funding agency (if they differ), but also the training program director, trainers, and trainees. Communication between the training and evaluation staff should be a planned, priority effort and not just happenstance. A prime task for the evaluator is to confer with all relevant groups. The evaluator is responsible for explaining what the evaluation is all about, eliciting, either formally or informally, the view and reactions of these groups, and maintaining open communication throughout the evaluation.

The third major influence of the evaluation is the availability of technical support. Given all the political and financial support imaginable, it is still necessary to have a competent evaluation staff to

design and carry out the effort in relation to the particular training program. The evaluator must be capable of gearing the evaluation to the program. The use of a set procedure for evaluation is as inappropriate as using a set training course for all occasions.

In addition to knowing what is an appropriate evaluation method, the evaluation staff also must be able to select an evaluation strategy within its own capability. In order to plan and carry out some of the more sophisticated techniques, the evaluation staff ideally should be familiar with the following areas: statistics (inferential and descriptive), design (experimental and quasi-experimental), evaluation literature, survey and interview methods, psychometrics (reliability, validity, scaling, etc.), and observational techniques. However, a simple, yet appropriate, evaluation can usually be conducted by the training, administrative, or personnel staff (with only limited technical assistance).

DEVELOPMENT OF CRITERIA

So far we have described evaluation in fairly general terms. We will now turn to some more specific issues that address how to decide if a training program has performed in the way that was anticipated.

As has been discussed, this decision is accomplished through a comparison between the goals initially set for the program and the results obtained. However, such a comparison is often difficult, since the goals are usually stated in a general way. What is needed is a set of standards that are precise enough and explicit enough so that the evaluator, or anyone else, can determine if the program has or has not met the goals.

Goals that satisfy these requirements are called "criteria." Mager (1975), in an influential book on the subject, lists some of the characteristics of useful criteria:

1. They are written in clear, specific language which is not easily misunderstood.
2. They describe the goal precisely.

3. They should be observable. Although the underlying process may be covert (e.g., attitude change or learning), there must be some performance or record of performance that indicates whether the goal has been met.
4. The conditions for observation of the performance should be stated.

An example might clarify Mager's point. In an installation evaluation, a training director needs to anticipate the support for a planned program. The support is the goal. The criterion is an explicit way of determining support. In order to develop this criterion, one needs to select: a method to gauge support; the individuals whose support is relevant to training; and the level of support deemed adequate in order to decide to implement the training program.

There are no universal answers to these questions. Each decision involved in developing criteria must be made in the light of practical concerns, appropriateness to the specific organization and setting, and technical adequacy. In the next sections, we will discuss the stages in the development of criteria: (a) the selection of the goals, (b) the methods used to evaluate the goals, (c) some of the standards to use in selecting a specific method, and (d) rules for deciding if the goal has been met. Since some of these topics have been discussed previously (Chapter X), some of our descriptions will be brief and we refer the reader to the previous chapter for added information.

The Selection of Goals

The development of criteria involves a process of specification and clarification. The first step in the process is the determination of the goals of the training course to be evaluated. Evaluations conducted at different stages in the development of a training program are concerned with different goals. We might note briefly some of the more important goals that are relevant at different stages.

In an installation evaluation, conducted prior to the development of the program, the evaluator might consider the degree that the proposed program fulfills the following goals:

(1) Will the course teach a skill, information, or an attitude that is needed?

(2) What proportion of the staff can use the course?

(3) How critical is the need for the course, relative to other training and to other organizational needs?

(4) Are there alternatives--cheaper and easier ways--to get the skill? Should selection of trained personnel or the use of already available programs be considered?

(5) Are there available resources--money, trainee time, trainer time, trainer skill and knowledge, etc.--to offer the course?

(6) Is the course acceptable to those involved?

(7) Can the skills or knowledge, once acquired, be applied in the organization?

When the training program is being developed and revised, using a formative evaluation strategy, the following points can be raised?

(8) Is the program content relevant to the skills, knowledge, or attitudes that are to be changed in the course?

(9) Is the training method an appropriate and effective one for teaching the skills, knowledge, or attitudes?

(10) Is the training presented in such a way that trainees attend to the material?

(11) Is the level of training consistent with the abilities and prior knowledge of the trainees?

(12) Can the training staff successfully present the course?

(13) Are trainees who take the course selected appropriately?

When the training course is fully developed, summative evaluation is used to judge its effectiveness. The issues to consider at this stage are:

(14) Is the program still needed?

(15) Do trainees find the course interesting, understandable, and useful?

(16) Have the trainees learned the content of the course?

(17) Do the trainees use the training back on the job, i.e., is the training transferable?

(18) Has the program affected organizational functioning?

(19) Has the program helped the agency carry out its goals?

(20) Are there any other behavioral or attitudinal consequences of training?

Selection of Standards

The foregoing are some of the more important goals of a training program. Often a set of goals must be selected from the list because of the time requirements and financial resources needed to carry out an evaluation of all of the goals. Even when all of the goals will be used in an evaluation, priorities must be made, since results relevant to different goals may be inconsistent, e.g., effective training may prove very costly, or training preferred by administrators may be uninteresting to trainees. In either case, standards need to be selected or ordered. The following are some rules we suggest using in selecting the goal or goals to be in an evaluation:

(1) The goals should be relevant to the type of evaluation (installation, formative, summative) being conducted.

(2) The goals should be relevant to the training. For example, if a course is mandated by the state government, questions about its organizational support or trainee approval seem less appropriate as goals than when a course is developed only to meet organizational needs.

(3) The goals should be measurable within a reasonable amount of time and for reasonable resource expenditure. For example, if one is installing a crisis intervention program in order to handle infrequent but disruptive problems, the goal of transfer (use of the learned behavior on the job) is not a reasonable goal.

(4) The relevant parties, those involved in making decisions about the funding, implementation, and use of training should agree that these are appropriate goals for the course.

(5) These goals have either been the most significant ones in making decisions about training in the past, or seem the most significant now. The goals used by decision makers to start or terminate a program need to be considered.

(6) If there have been problems with some aspect of prior training, goals relevant to this aspect should be evaluated. For example, if prior training programs have had problems because trainees are bored, then trainee interest needs to be used as a goal.

(7) The most central goals of the course, i.e., what it tries to teach, should be evaluated.

Method

After selecting the goals to use in the evaluation, one needs to decide which methods can be used to determine if the goals have been met. We list some of the methods that can be used in evaluating each of the training goals we discussed before. Table XI-1 summarizes the methods that can be used in evaluating all of these goals.

Program Installation Evaluation

For a program installation evaluation, the following goals might be considered:

(1) Course teaches a needed skill, information, or attitude.

(2) High proportion of the staff can use the course.

(3) Course is of considerable significance, relative to other training and to other organization needs.

(4) Alternative (cheaper and/or easier) ways to get the skill.

(5) Resources--money, trainee time, trainer time, trainer skill and knowledge, etc.--are available to offer the course.

(6) Course is acceptable to those involved.

Table XI-1

Methods That Can be Used to Evaluate Training Goals

Training Goals	Interviews with Staff (a)	Questionnaires to Staff (b)	Observation of Job (c)	Review Organization Records (d)	Review Laws & Professional Standards (e)	Review Performance Standards (f)	Content Validity Methodology (g)	Interviews with Trainees (h)	Tests of Knowledge and Skill (i)	Measures of Trainee Interest (j)	Observation of Training Session (k)	Interviews re Training Techniques (l)	Records of Prior Training (m)	Observation of Performance (n)	Interviews with Supervisors (o)	Questionnaires to Trainees & Supervisors (p)	Contact Other Agencies & Persons (q)
Program Installation Evaluation																	
1. Training needed skill	X	X	X	X	X	X											
2. Proportion of staff has need	X	X	X	X	X	X											
3. Significance of need	X	X	X	X	X	X											
4. Alternatives available	X																X
5. Available resources																	X
6. Course acceptance	X	X	X														
7. Apply skills on job	X	X	X														
Formative Evaluation																	
8. Content relevant to skills	X		X				X										
9. Method appropriate											X	X					X
10. Promotes interest								X	X	X	X	X	X				
11. Appropriate level								X	X	X	X	X	X				
12. Staff able to teach								X	X	X	X	X	X				
13. Selection of trainees								X	X	X	X	X	X				
Summative Evaluation																	
14. Program still needed	X	X	X	X	X	X											
15. Course is interesting								X		X						X	
16. Learning									X					X			
17. Use training on job														X	X	X	
18. Affect organization														X	X	X	
19. Affect organization's goal	X			X													
20. Other consequences														X	X	X	

(7) The skills or knowledge acquired can be applied in the organization.

The methods appropriate for evaluating Goals 1, 2, and 3 are:

(a) Interviews with administrators, staff, and supervisors.

(b) Questionnaires given to staff and supervisors.

(c) Observation of the job behavior.

(d) Review of organization records.

(e) Review of laws and professional standards.

(f) Review of performance standards developed for workers.

Goals 6 (Acceptance) and 7 (Applicability) can be evaluated using the first three methods above. Goal 5 (Available Resources) can be evaluated, in part, by writing correctional agencies or colleges near one's agency or CONtact, Inc. to get information about other available training. If selection is considered a potential alternative to training (Goal 4), the evaluator should contact the agency's personnel staff to discuss the availability of individuals with the needed training and the probable salary that would be required to attract such individuals.

Formative Evaluation

In a formative evaluation, used to evaluate and revise a training program which is being developed, the following goals are considered potentially useful in an evaluation:

(8) Program content is relevant to the skills, knowledge, or attitudes that are taught in the course.

(9) Training method is an appropriate and effective one for teaching the skills, knowledge, or attitudes.

(10) Training is presented in such a way that trainees attend to the materials.

(11) Level of training is consistent with the abilities and prior knowledge of the trainees.

(12) Training staff can successfully present the course.

(13) Trainees who take the course are appropriately selected.

Content validation methodology (g), should be used to evaluate Goal 8. Content validation involves the logical comparison of the course content--lectures, discussion topics, course assignments, practice problems--with the performance goals of the course (as determined by a thorough job and task analysis).

Goals 10 (Attention), 11 (Consistent with Ability of Trainees), 12 (Staff Ability), and 13 (Selection of Trainees) can be evaluated using the following methods:

(h) Interviews with trainees.

(i) Tests of knowledge.

(j) Measures of trainee interest.

(k) Observation of training sessions.

(l) Interviews with the training staff and trainees to determine which training techniques have been most effective in the past.

(m) Examination of evaluations or records of prior training.

Goal 9 (Training Method is Appropriate) is one of the more difficult goals to evaluate. Often there is no information available to determine if a training method will be effective for a certain content area and for certain types of trainees. Also, some of the relevant information available in research reports and published articles may be very technical and hard to understand. For the evaluator who cannot locate or prefers to avoid the use of research on the training method being used, methods (l) and (m) seem reasonable to use to evaluate Goal 9.

Summative Evaluation

In a summative evaluation, used for evaluating the effectiveness of a completely developed course, the following goals may be used:

(14) The program skill is needed.

(15) Trainees find the course interesting, understandable, and useful.

(16) Trainees have learned the content of the course.

(17) The knowledge, skills, and attitudes acquired in the course are used back on the job.

(18) The program has affected organization functioning.

(19) The program has helped the agency to carry out its goals.

(20) There are other behavioral or attitudinal consequences of training.

Goal 14 (Skill is Needed) can be evaluated using the same methods as used for Goal 1, described previously. Goal 15 (Interest) can be evaluated using interviews or surveys of trainee interest. Goal 16 (Learning) can be evaluated using a test. Goal 17 (Transfer), 18 (Organization) and 20 (Other Consequences) can be evaluated in the following ways:

(n) Observation of employee performance.

(o) Interviews with supervisors of trainees.

(p) Questionnaires given to trainees or supervisors.

Goal 19 (Agency Goals) is often difficult to evaluate, since the relevant events may be infrequent, e.g., crisis situations, or may be difficult to measure, e.g., improved client functioning. The most frequently used and accessible measure of this goal involves the use of agency records, e.g., recidivism rates, or percentage of clients who have jobs six months after release.

Selection of Methods

Thus far, we have discussed: the goals that can be used in an evaluation of training; the selection of specific goals for the evaluation; and the methods that can be used to measure the achievement of these goals. The next step in the development of criteria involves the selection of a specific measurement method or methods. We suggest considering the following issues in selecting a method:

(1) Expense. In selecting a method, the evaluation needs to consider the cost of the method, e.g., the questionnaires, and the time of staff involved in using the method.

(2) Cooperation. One of the major problems with questionnaires and any other frequently used measurement method is the difficulty in eliciting respondent cooperation. Respondent's honesty and help are crucial to performing a useful evaluation.

(3) Ability. Do the evaluators know how to use the measurement technique and interpret its results?

(4) Understandability. Can the persons who will make decisions about training understand the results of the evaluation using the method?

(5) Validity. To what degree does training determine scores on the measuring instrument? In other words, are there other significant influences on the measure. One of the major concerns in using measures of organizational changes, and offender behavior as indicators of training effectiveness is the degree to which many other factors may influence these measures. What is desired is a measure that is primarily affected by the training.

(6) Acceptance. Do the decision makers accept the method as appropriate and effective?

(7) Availability. If the evaluator is considering the use of records, e.g., turnover, recidivism, are the records available?

(8) Bias. Bias occurs when characteristics of the measurement process leads to inaccurate inferences (Anderson, Ball, Murphy, & Associates, 1975). The common sources of bias in evaluating training are:

(a) Social desirability. People are often reticent to give negative responses on questionnaires or interviews and may distort records in order to present the organization and their own behavior in a desirable light.

(b) Regression effects. There is a statistical artifact which may occur if pre-training measures and post-training measures are used. If individuals who score either very high or very low on the

initial measure are selected as the only participants in the evaluation study, the results of the research may be misleading. Extreme scores tend upon retesting to become less extreme (i.e., improve) so that statistical artifacts may be misinterpreted as training effects.

(c) Reactivity. If individuals know they are being observed, they may act differently (e.g., "Hawthorne" Effect). For example, trainees may feel that the organization is especially concerned about their performance and they may be very careful in their behavior.

(d) Changes in the measurement. Certain measures may change over time. For example, interviewers may become more skillful or records may become more or less accurate. Comparison of measures before and after training may be influenced by these inadvertent changes.

(e) Retesting effects. If pre-training and post-training measures are used in evaluation, trainees may become sensitive to the issues on the pre-training test and perform better on the post-test independently of training.

Decision Rule

Having selected a specific method to use in evaluating each goal, the next step is to determine a decision rule. The decision rule concerns the degree of change that is accepted as evidence that training has met its goal. For example, if an evaluator found a 10 percent improvement in job performance ratings, or a 15 point increase in test scores, is this enough of an improvement to conclude that training was effective? Our recommendation is that the evaluator rely less on statistics, e.g., tests of statistical significance, and more on organizational agreement as decision rules.

At the beginning of the evaluation effort, the individuals involved in training decisions must select the minimum change that is acceptable. In making this decision, several issues must be considered--cost-effectiveness, reasonableness, and the type of goal. Cost-effectiveness concerns a comparison between the costs of training and its effects. It is

reasonable that costly programs either train more people or train people more effectively. Cost-effectiveness has limits, however. The limits are those of reasonable expectations. Training, especially short-term training, cannot be expected to radically revise behavior. If the decision rule is unrealistically high, training will almost always fail. A final issue to consider is the nature of the goal. The more removed the goal is from the training, the less impact the training is likely to have. For example, a very small change in recidivism is a reasonable goal. There are too many factors that influence offender behavior than to expect training to have any but a small impact.

Statistical tests can also be used in reaching this type of decision. Such tests can be used to determine the significance of the differences due to training. Some of the relevant statistics are described in the last part of this chapter. Although statistical tests can facilitate decision making, they are not essential for making decisions about training, and if used without sufficient knowledge of their meaning, they can lead to incorrect inferences about training. Regardless of the use of statistics, organizational discussion of and agreement about expected changes that result from training are essential in the evaluation. All those involved in training need to agree on what changes are adequate to maintain the course. In addition, reading about changes resulting from other training may not only give those involved a realistic set of expectations, but also knowledge about other types of training and evaluation.

To summarize, we recommend that the decision rule regarding the impact of training should be acceptable to those involved, be relevant to the cost of the program, and be realistic considering the nature of the goal.

Illustration of Criterion Composition

In the beginning of this section, we stated that a criterion had three components: a goal; a method; and a decision rule. Here are two examples of the components of a criterion:

Example 1:	Goal 17:	Trainees use the communication skills acquired in training on the job.
	Method:	Interviews with supervisors before and three months after training about the trainee's use of certain specific communication skills.
	Decision Rule:	If supervisors report an average increase of 25% in the use of the skills, the training is judged as meeting its goal.
Example 2:	Goal 11:	The level of training is consistent with the abilities and knowledge of the trainees.
	Method:	Tests of trainee's knowledge before training and questionnaires given during training.
	Decision Rule:	If the average grade on the pre-training exam is 50% or less, we judge that trainees do not already know the course content. If 85% of the trainees indicate on the questionnaire that the lectures and reading materials are understandable, the level of the course is judged appropriate to the trainee's ability. If both conditions are satisfied, then the goal has been met.

EVALUATION DESIGN AND IMPLEMENTATION

After selecting a set of criteria, the next step in the evaluation process involves determining the way the evaluation will be conducted. One reason the evaluation design is of special importance is that an appropriate design may reduce the impact of bias (distortion of results) in the evaluation.

There are a number of designs that are appropriate to use in an evaluation of training (see reviews of Cook & Campbell, 1976; Rossi, Freeman, & Wright, 1979). We do not think that we can adequately and briefly explain some of the more involved designs. Nor do we believe that such designs are essential for an effective evaluation. We propose to describe some simple designs that can be carried out by individuals without extensive knowledge of statistics and experimental design. If a more sophisticated approach is deemed necessary, then the evaluator should refer to experimental design and evaluation books, e.g., Campbell & Stanley, 1963; Cochran & Cox, 1957; Kirk, 1969; or Myers, 1972.

Designs Using Control Groups

Control groups are groups used for comparison. If a researcher wants to determine the impact of a training course, the researcher can compare the performance of individuals after training with the performance of the control group--individuals who do not receive training or who have received a different type of training.

The control group is selected to be as similar to the group receiving the training (the experimental group) as is possible. The best procedure for selecting members of an experimental and a control group is through random assignment. Random assignment is a process wherein all the people in the study have an essentially equal chance of being in the experimental and the control group.[1] If the total number of people in the sample is 60 or more, random assignment tends to equalize the groups.

In most training evaluation studies there are fewer than 60 participants. Under such conditions, random assignment is unlikely to equalize the groups. The alternative procedure is to match individuals in the

[1] For example, toss a die for the assignment of each person: if the die is odd, the person is in one group; if the die is even, the person is in the other group. Or use the last digit of the social security numbers for odd-even assignment.

control and the experimental groups. The groups need to be matched on variables that should influence training results. For a training study, individuals in the groups should be matched on prior experience or knowledge about the course content, ability, and education.

Matching is more problematic than random assignment. It may not be possible to get information about the relevant variables to be matched, and it also may be difficult to find individuals who are equal on all of the relevant variables.

Regardless of which method, random assignment or matching (even if less than perfect), is used to equalize the experimental and control groups, the use of control groups improves the quality of the evaluation research. It can be used to reduce the effects of the bias problems previously discussed. If a control group is used, any differences in performance between the experimental and control group is likely due to training.

Is the use of a control group feasible? In many instances, the number of participants in a training course must be limited. The restriction of size is especially appropriate when a course is new. When the course is limited to fewer employees than need and/or want the course, those not receiving training (this time) might constitute the control group. Even if such restriction is not absolutely essential, use of size limitation in order to have a control group is worth the extra time, effort, and expense.

Pretest-Posttest Designs

The purpose of random assignment and matching is to equate the experimental and control groups. Pretests can be used as an additional guarantee that the groups are, indeed, equated. If the groups are not equal on the pretest, a statistical procedure, covariance analysis, can be used to determine the impact of training as distinct from any initial differences in ability or knowledge in the groups.

Often pretests are used without control groups. Training effects are equated with any differences between pretest and posttest performance in the training group. This design, however, is open to a number of biases--changes may result from exposure to the correctional environment and agency personnel rather than training; there may be retesting effects, or changes in measurement (both discussed in reference to selection of methods). Thus, this design is not recommended.

Time Series Designs

There may be certain situations when it is not possible to have a control group, e.g., when all available individuals must take the training course. Under these circumstances, a revised version of the pretest-posttest design is appropriate. In this design, pretest measures of performance, either through observations, interviewers, or tests are given several times before training. Below is a diagram of this design:

$Pretest_1$---->$Pretest_2$---->$Pretest_3$---->Training---->$Posttest_1$

---->$Posttest_2$---->$Posttest_3$

This design handles many of the problems of bias. In addition, the stability of the changes due to training can be assessed. However, there are certain difficulties associated with this method. If the measurement technique involves interviews or questionnaires, repeated pretests and posttests will result in reduced cooperation, if not surliness and physical abuse. This is much less a problem, if workers are only observed, especially if the observation is low-key. An additional problem is the expense involved in using this design.

Combination Designs

It may not be feasible to use an appropriate evaluation design. Should one then forget evaluation? If the situation allow no reasonable evaluation, it is probably better to refrain from performing a poor evaluation that may lead to distorted conclusions. There is, however, an alternative option. Several designs, none of which alone can evaluate training, together may be effective. With each study, one's evidence about the course is increased.

One could begin by waiting several months after training and comparing performance records, e.g., performance appraisal or records of disciplinary problems, percentage of employees dismissed, and similar records of employees in the same job last year who had about the same amount of experience. Information about the education, selection test scores, if any, and prior experience of the two sets of employees would help one to determine how adequate last year's sample is as a control. If the groups are reasonably comparable, then the differences in ratings are more likely due to training.

A major problem in such a study is the possible role of specific events during or after the training that influenced the results. Perhaps, an incident caused some changes in agency policy leading to different employee behavior. The experimental group's performance might be improved due to this policy and not the training. To assess this, one needs to find another sample of employees as similar as possible to those in the experimental group. Those individuals (which can be considered to a quasi control group) would be given pretests and posttests at the same time as the trainees.

We are not trying to suggest that any evaluation design, no matter how open to bias, should be used. We are saying that several such designs, <u>together</u>, may make up a reasonable evaluation study. If only one such limited design is possible, we recommend postponing the evaluation until better evaluation designs can be used.

DATA ANALYSIS

The last step in an evaluation involves analyzing the results. The results are used, along with the decision rule, to decide if the training has been effective. The decision rule, as described before, is a standard of the amount of change expected from training. If the results meet or exceed the standard, the training is judged as effective. For example, individuals involved in training might decide that a 10 percent reduction in assaults involving offenders is expected from an extensive training program in crisis intervention. If the training results in a 12

percent reduction, the training can be said to have met its goal. If, however, there is only a 5 percent reduction, the training is judged as not meeting its goal.

If decision rules are based on judgments, one needs a way to summarize the results so they can be compared to the standard. The method used for summarizing results is called descriptive statistics. Examples of these are frequencies, percentages, means (averages), and mean differences.

Some descriptive statistics that might be used for analyzing training data include:

- The average score on a posttest training exam (as compared with the average score on the pretest exam).
- The difference between the average test scores of the experimental groups and the control groups.
- The frequency with which a certain behavior is performed after training.
- The difference in frequencies of behavior before and after training.

As discussed previously, inferential statistics can also be used as decision rules. When using inferential statistics, the interest is on the likelihood that the performance change after training might be due to chance events, rather than to training. In other words, if the evaluation were repeated, would similar results be obtained?

There are numerous statistics texts which can be used as an aid in analyzing data. We refer you to these texts to learn the procedures involved in carrying out statistical tests, since description of these tests is beyond the scope of this report. Some of the frequently used statistical tests that can be used with some of the analyses previously discussed are:

Design	Statistical Test
Posttest only, with random assignment into experimental and control groups.	Mean difference tests, e.g., t-test
Posttest only, with matching of the experimental and control groups.	Matched group t-test
Pretest-Posttest design.	Matched group t-test
Time series design.	Analysis of covariance (comparing pretest and posttest groups)

A FINAL NOTE

We have outlined many conditions that should govern in an evaluation. Rarely will all conditions be met. But we urge the evaluator to try assiduously to meet as many as possible, to be relatively hard-nosed and inflexible in evaluation demands, to be honest and "pull no punches" in reporting evaluation results. What we propose is this: (a) do conduct evaluations, (b) do not withhold evaluating because you do not have good evaluation data--a "best as you can" evaluation is likely to have more effect on the agency than no evaluation at all. The better your evaluation procedures, the better your programs are likely to become, or the more efficient and effective your organization is likely to be. We have offered caveats with respect to the requirement for rigid experimental design and statistical analysis. But one should recognize that sound statistical and design procedures are likely to be necessary if one intends to publish evaluation results in most professional journals.

CHAPTER XII

PHASE II: PLAN FOR DEMONSTRATING THE EVALUATION MODELS

OBJECTIVE OF PHASE II

The objectives of Phase II of the project are to:

(1) Develop assessment and evaluation models with appropriate evaluative criteria that can be applied across a wide range of correctional personnel training programs in a wide range of training environments.

(2) Demonstrate the usefulness of the models by applying them to three training programs selected especially to test and illustrate the broad applicability of the models.

Phase I has produced information (through literature review, site visits, and a national survey of correctional personnel training) to influence the ultimate forms of the models. The Phase II demonstration should show how evaluation concepts, practices, and procedures that have been formulated can locate, define, and indicate corrective action for differences between expected job performance and actual job performance.

BACKGROUND

Model Development

Phase I of the study of training of correctional personnel has produced information on the many variations in content of training, training methodologies, and other factors influencing training. In Phase I we developed (a) a very complete conceptual, generic. Instructional System Operations Model of the training process; (b) an Instructional System Evaluation Model of the evaluation process, both for formative evaluation (evaluation of the training process itself), and for summative evaluation (evaluation of the outcome of a training program); plus (c) a Correctional Issues Model that illuminates the valuative, policy, and practice issues in correctional personnel training. These three models have been tailored specifically to training of correctional personnel using the knowledge

gained during Phase I site visits and that acquired through the national survey. Their utility will be tested and demonstrated in Phase II of the project.

We have devoted considerable attention to issues and policies in the field of corrections that often have an overriding influence on training. It has become clear to us that the issues that evolve in the Correctional Issues Model--the turbulent nature of the corrections environment, the organizational climate, individual worker dissatisfactions with progress, change, and reward in corrections work, and changing the attitude of correctional job incumbents and correctional organizations from survival in the correctional environments to innovation and problem solving to effect change and improvement in the system--are and should be explicit considerations in both of the other two models.

Both the Instructional System Operations Model and the Instructional System Evaluation Model are deductive logic models of the complete training and evaluation process. We have used the Operations Model as a general, all-inclusive guide to assess the development and completeness of training. As we have reported earlier, it is rare that training programs are complete with respect to all aspects of the model. Although this does not diminish the usefulness of the model, the model will be improved as it is able to emphasize those parts of programs that appear to suffer in their development or implementation because of lapses or inattention to particular aspects of model processes. Site visits have suggested where potential deficiencies are likely to occur; survey results help corroborate and elaborate site visit findings.

Each of the three models requires a different perspective in applying the model to a specific training program. Each model presents a full set of requirements that must be addressed and the concerns that must be taken into account in examining a training program. Questions that must be answered are laid out for each model. The extent to which the questions can be answered becomes the criterion by which to judge the thoroughness of an evaluation design.

Current Correctional Personnel Training Departs from Models

Phase I assessments of training programs show some fairly general lapses in program development and evaluation--in job analysis, in establishment of training needs, in assembling the training program, in conducting the program, in determining the outcome of training.

Need for Training not Established

Training establishments must become aware that sound training for work cannot be developed unless it is known what kind of worker performance is required to do that work satisfactorily, and then to be able to demonstrate by way of performance data that desired performance is not being achieved. The more evident aspects of front-end analysis often are neglected:

- The job and work place should be defined and described, with the product(s) to be expected if the job were done properly.
- There should be an understanding of the nature of the job in the context of the whole correctional organization.
- Physical, mental, temperamental, personality, and attitudinal requirements of workers on the job should be established.
- Standards of worker performance expected for each task of the job should be established.
- The social, political, and legal constraints that impact on the job itself and on worker outputs ought to be known.
- One should ask whether training is the most cost-effective way to correct a recognized performance gap, or whether administrative/organizational/policy decisions might do it more easily and better.
- The importance of performance measures and records to define and assess work should be recognized.

To the extent possible, job analysis, performance standards, and performance measurement need to be objective, and stated in quantitative terms capable of being scaled or ranked.

Training Program Development and Implementation not Complete

The most significant departures from the operations model deal with the training program itself, with lack of attention to: (a) defining the knowledge, skill, ability, and attitude changes that training should bring about, and (b) selecting the most appropriate way to bring about those changes.

The knowledge, skill, ability, and attitudes required for a job should be specified in the job and task analyses. The extent to which those attributes are lacking in the population of potential trainees often is conjectural. Furthermore, the relation of an attribute to performance is likely to be low, so even if there were a perfect match of lack of attribute with emphasis of that attribute in training, the effect on performance usually will be less than hoped for. The point made is simply this: to make training optimally effective, the pertinent characteristics of the job incumbent need to be assessed accurately, accurate estimates need to be made of these attributes in task analysis/job analysis/performance standards, and the training program then designed to match the deficiencies exposed by job performance.

Outcome of Training not Assessed

There has been little rigorous investigation of the outcome of training programs. Many programs exercise the pro forma process of "before" and "after" testing to learn what was liked or disliked, to learn if expectations were met, if attitudes, perceptions, and confidence in one's skills have changed. But the more difficult tasks of comparing performance on the job before and after training, assessing transfer of training, is not done. More careful observation and study of the long-term effects of training are needed. The critical elements, usually missing, are objective performance measures, the same measures considered under the heading of performance standards.

These criterion measures must be both reliable (consistent, capable of being reproduced) and valid. Validity should be demonstrated. It should be shown that the measure truly and accurately represents an

important part of a worker's performance. Such measures are likely to be factors that have been identified in the job and task analysis, as well as in the performance standards.

For example, consider the line correctional officer whose principal responsibility is likely to be seen as maintaining security within the institution, probably with minimum use of physical force. In the broadest sense, security can be measured by the number (per unit of time) of escapes, or of major (defined) riots or disturbances, of institution custodial personnel assaulted, injured, or killed. But these are very broad measures--the numbers rarely will exhibit enough range to demonstrate validity--especially over a short period of time. There are likely to be many corollary, more proximate measures that relate to security in the broadest sense, but will assess it in more narrow terms: number of incidents in which any kind of physical restraining force needed to be used; number of grievances or number of law suits filed by inmates; number of requests (made to the correctional officer) for advice, counsel, or assistance; accuracy and quality of the correctional officer's record-keeping; performance ratings of correctional officers by supervisors; peer ratings among correctional officers as to which other officers the rater would like to work with on a shift. Also, most measures will differ widely in scale as a function of the nature of the institution (a high security institution, as contrasted with one less so), the duty shift to which they apply (more opportunity for disturbances when inmate interaction is greatest--e.g., meal time, exercise time, mail call, medicine distribution, etc.), and so on.

DEMONSTRATION OF THE MODELS

We propose a demonstration of the evaluation models in three correctional agencies by way of three separate 3-day workshops.

Selection of Sites for Demonstration

We plan to select three correctional agencies where we can demonstrate the evaluation models. In order to best test the models, the agencies chosen should display functional, geographic, and size diversity. Although not critical to the demonstrations, a prison, a probation and

parole agency, a pre-release center, or a juvenile detention agency would offer the appropriate kind of functional and size diversity. The South, Midwest, and East regions of the United States will provide sufficient geographic differences among the demonstration agencies.

We will ask that each agency commit its training personnel staff and related administrative staff to participate in a 2- or 3-day workshop in February or March 1980. The agencies selected should have no training scheduled for the 3 days of the workshop and should not have any significant number of its training staff on leave, or holiday, or otherwise away from the organization. There should be a classroom of sufficient size available for the workshop.

Conduct of Workshops

Three ARRO project staff members will conduct the workshops. For all the classroom portions of the workshop, we plan to use a modified "team-teaching" technique. This instructional method was observed during site visits to California and Colorado to be very effective in teaching crisis intervention, as well as other topics. The method requires that two instructors be equally well trained to handle the presentation. They often are in front of the class, concurrently, each holding the floor no longer than a few minutes at a stretch, before turning to the other instructor for comment, or being interrupted by the other instructor with comments or suggestions. Both trainers must be thoroughly familiar with the subject matter, each must recognize his or her instructional and informational strengths and weaknesses, and both must have the common interest of keeping the class members motivated and actively participating in the training.

It is our intention to encourage much discussion and give-and-take through the course of the workshop. As will be described, workshop participants will be given active roles in evaluation. Through the participation, all should become involved in looking critically at their own work. The workshop culminates in asking participants to examine critically, as well, the evaluation models we have promulgated, and at us, as trainers.

Content of Workshop

In the schedule that follows, training is proposed to require 3 days. If an agency cannot commit its staff for 3 days, we are able to collapse the training into 2 days.

Day 1--Morning

Objectives of workshop.

Why evaluate?

When to evaluate.

Description of evaluation models (with flow charts and complete sets of evaluation questions for each).

- Instructional System Operations Model
- Instructional System Evaluation Model
- Correctional Issues Model

Distribution to participants of Chapters I, III, X, and XI of the Phase I report.

Day 1--Afternoon

Continue discussion of models.

Determine which agency training program(s) on which to focus evaluation models.

Divide workshop into three sections, one section for each model; assign each section responsibility for examining the agency training program in the context of that model.

Assign an ARRO project staff member to work with each section.

Sections assemble as sections; each group nominates a leader for that section.

Sections spend remainder of afternoon finding answers to and documenting answers to questions posed by the model for which each section has responsibility; in subsequent Day 2 and Day 3 sessions, the section having responsibility for a particular model takes the lead in discussing that model.

Day 1--Evening

No-host social hour.

Day 2--Morning

All workshop participants--discussion of the Instructional System Operations Model as it applies to the agency training program selected--although the session will be led by ARRO project members, the section assigned that model will take the lead in discussion of the questions the model poses.

Day 2--Afternoon

All workshop participants--discussion of the Instructional System Evaluation Model as it applies to the agency training program selected, with the same instructional concept as for the previous session.

Day 3--Morning

All workshop participants--discussion of the Correctional Issues Model as it applies to the agency training program selected, again, with the same instructional concept as before.

Day 3--Afternoon

All workshop participants--apply the three models in evaluating the workshop itself as a training program--evaluating both the workshop content, and the ARRO project staff as trainers.

Obtain commitment of workshop members to return a comment questionnaire (about the workshop and the impact that such evaluation models have on their programs) sent to them a week after the workshop is concluded.

We believe the workshop format is a good way of presenting the requirements for program evaluation. The entire training staff becomes involved in all aspects of an evaluation in the context of examining their own programs. We hope a greater appreciation of the evaluation process and the benefits of evaluation will develop.

Assessing Use of Evaluation Materials

Evaluation models ought to be used. The workshop will encourage use of the models. It will be important that LEAA continue association with the three corrections agencies that participate in the demonstrations to learn if familiarity with evaluation processes has had any effect on agencies' programs.

The Executive Summary of Phase I is to be distributed widely among corrections agencies. It should be expected that agencies that have the greatest interest will request the full report. As with the three workshop agencies, it will be desirable to learn if these requestors of the report ever use the information therein. Plans for the continuing assessment, both of the demonstration agencies and the report recipients, will be elaborated in the Phase II report.

PHASE II REPORT

An account of the workshop proceedings will be prepared, including:

- Perception of the interest and enthusiasm for evaluation in general.
- Applicability of the models to the programs evaluated.
- Changes in the models that occurred as a consequence of the demonstrations.
- A proposed strategy for follow-up assessment of evaluation activities at the three workshop locations 6 to 12 months after the demonstration sessions.
- A proposed strategy for implementing and monitoring, at one or more locations, evaluations using the models developed in the project.

REFERENCES TO FULL REPORT

Adams, W. J., Tabor, L. D., Baker, C. E., & O'Neil, M. Community resources training center for corrections--final report. Olympia: Training Section, Division of Social Institutions, Department of Social and Health Services, November 1970.

Alden, J. Evaluation in focus. Training & Development Journal, 1978, 32, 46-50.

Aldrich, H. & Herker, D. Boundary spanning roles and organization structure. Academy of Management Review, April 1977, 217-230.

American Correctional Association. Correction officers' training guide. College Park: Committee on Personnel Standards & Training, 1978.

American Correctional Association. Juvenile and adult correctional departments, institutions, agencies, and paroling authorities, Directory, United States and Canada, 1979 Edition. College Park, 1979.

Anastasi, A. Psychological testing (4th Ed.). New York: MacMillan, 1976.

Anderson, S. B. & Ball, S. The profession and practice of program evaluation. San Francisco: Jossey-Bass, 1978.

Anderson, S. B., Ball, S., Murphy, R. T., & Associates. Encyclopedia of educational evaluation. San Francisco: Jossey-Bass, 1975.

Baker, E. L. Lean data strategies for formative evaluation. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 1976.

Bilek, A. J. America's criminal justice system--A diagnosis and prognosis. In Criminal Justice Monograph: The change process in criminal justice. National Institute of Law Enforcement and Criminal Justice, U.S. Department of Justice, June 1973.

Blum, L. N. Evaluation of correctional training programs. Training and Development Journal, 1976, 243-350.

Brown, W. B. Systems, boundaries, and information flow. Academy of Management Journal, 1966, 9, 318-327.

Brown, B. S. & Sisson, J. W., Jr. The training program as a correctional change agent. Crime and Delinquency, July 1971, 302-309.

Brutvan, D. R. Final report for the planning-development and first year operations phase of the Erie County staff training and inmate rehabilitation project. Buffalo: State University of New York, Division of Continuing Education, June 1969.

Bunker, K. A. & Cohen, S. L. Evaluating organizational training efforts: Is ignorance really bliss? Training and Development Journal, August 1978, 4-11.

Burkhart, B. R., Behles, M. W., & Stump-Hauzer, J. S. Training juvenile probation officers in behavior modification: Knowledge, attitude change, or behavioral competence? Behavior Therapy, 1976, 7, 47-53.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Campbell, D. T. A study of leadership among submarine officers. Columbus: Personnel Research Branch, Ohio State University, 1953.

Carter, R. M. The diversion of offenders. Federal Probation, March 1972, 31-36.

Catalanello, R. F. & Kirkpatrick, D. L. Evaluating training programs--The state of the art. Training and Development Journal, May 1968, 2-9.

Clemmer, D. The prison community. New York: Rhinehart, 1958.

Cochran, W. G. & Cox, G. M. Experimental designs. New York: Wiley, 1957.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for adult parole authorities. Rockville: July 1976.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for adult community residential services. Rockville: April 1977a.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for adult probation and parole field services. Rockville: July 1977b.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for adult correctional institutions. Rockville: August 1977c.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for adult local detention facilities. Rockville: December 1977d.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for juvenile community residential services. Rockville: April 1978a.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for juvenile probation and aftercare services. Rockville: July 1978b.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for juvenile detention facilities and services. Rockville: February 1979a.

Commission on Accreditation for Corrections, American Correctional Association. Manual of standards for juvenile training schools and services. Rockville: March 1979b.

Cook, T. D. & Campbell, D. T. The design and conduct of quasi-experiments and true experiments in field settings. In M. D. Dunnett (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

Darling, L. Still Liddy after all these years. Stranger than fiction: The prison exploits of G. Gordon Liddy, and other eerie tales. Washington Post, October 27, 1979.

Dewey, J. The public and its problems. Denver: A. Swallow, 1927.

Dill, W. R. Environment as an influence on managerial autonomy. Administrative Science Quarterly, 1958, 2, 409-443.

Dinitz, S. & Beran, N. Community mental health as a boundaryless and boundary-busting system. Journal of Health and Social Behavior, 1971, 12, 99-108.

Dorin, D. & Johnson, R. The premature dragon: George Jackson as a model for the new militant inmate. Contemporary Crises, 1979, 3, 295-315.

Dunnette, M. D. Personnel selection and placement. Monterey: Brooks/Cole, 1966.

Edwards, W., Guttentag, M. & Snapper, K. A decision-theoretic approach to evaluation research. In E. L. Streuning and M. Guttentag (Eds.), Handbook of evaluation research. Sage Publications, 1975.

Evan, W. M. The organization-set: Toward a theory of interorganizational relations. In J. D. Thompson (Ed.), Approaches to organizational design. Pittsburgh: University of Pittsburgh Press, 1966.

Farmer, R. E. Cynicism: A factor in corrections work. Journal of Criminal Justice, 1977, 5, 237-246.

Fine, S. A. Personal communication, October 1979.

Fine, S. A. & Wiley, W. W. An introduction to functional job analysis: A scaling of selected tasks from the social welfare field. Washington, D. C.: W. E. Upjohn Institute for Employment Research, 1971.

Fogel, D. The politics of corrections. Federal Probation, 1977, 41, 27-31.

Frank, B. The emerging professionalism of the correctional officer. Crime and Delinquency, 1966, 12, 272-276.

Gilman, M. Problems and progress in staff training. Crime and Delinquency, 1966, 12, 254-260.

Gilson, J. W., Hagedorn, J. D., & Crosby, R. F. Regionalized personnel development for criminal justice employees. Journal of Criminal Justice, 1974, 2, 243-248.

Glaser, D. The new correctional era: Implications for manpower and training. Crime and Delinquency, 1966, 12, 209-216

Glaser, D. Politization of prisoners: A new challenge to American penology. American Journal of Corrections, 1971, 33, 6-9.

Glickman, A. S. Effects of negatively skewed ratings on motivation of the rated. Personnel Psychology, 1955, 8, 39-47.

Goldstein, I. L. Training: Program development and evaluation. Belmont: Wadsworth, 1974.

Greenberg, D. F. Problems in community corrections. Issues in Criminology, 1975, 10, 1-29.

Grenfell, J. Training for corrections: Rationale and techniques. Carbondale: Southern Illinois University, Center for the Study of Crime, Delinquency, and Corrections, 1967.

Grzegorek, A. E. & Kagan, N. A study of the meaning of self-awareness in correctional counselor training. Criminal Justice and Behavior, 1974, 1, 99-122.

Hamblin, A. C. Evaluation and control of training. London: McGraw-Hill, 1974.

Hanford, S.,& Moone-Jochums, S. Evaluation model for training programs. Training Resource Center, Illinois Corrections Training Academy, 1978.

Harris, P. R.,& Harris, D. L. Increasing organizational effectiveness of military human resources: Action research in correctional training. La Jolla: Management & Organization Development, Inc., 1972.

Hasenfeld, Y. People processing organizations: An exchange approach. American Sociological Review, 1972, 37, 256-263.

Herzog, F., Trenholm, S., & Whiteneck, G. Crisis intervention training: An experimental evaluation program in immediate action decision making. (LEAA project #72-C3-(2)-C), February 1974.

Horne, A. N., & Passmore, J. L. Inservice training in a correctional setting: Facilitating change. Federal Probation, June, 1977.

Hosford, R. E., George, G. O., Moss, C. S., & Urban, V. E. The effects of behavioral counseling training on correctional staff. Teaching of Psychology, 1975, 2, 124-127.

Jacobs, J. B. What prison guards think: A profile of the Illinois force. Crime and Delinquency, 1978, 24, 185-196.

Jansyn, L. R., Jr. Correctional officer training in American correctional institutions. In Training the correctional officer, proceedings of two workshops. Carbondale: Southern Illinois University, Center for the Study of Crime, Delinquency, and Corrections, 1965.

Jensen, D. R., Schwartz, I. H., & Rowan, J. R. Evaluation of staff training programs. Wisconsin Division of Corrections, Department of Health and Social Services, John Howard Association, 1975.

Jesness, C. F., Allison, T. S., McCormick, P. M., Wedge, R. F., & Young, M. L. The cooperative behavior demonstration project. Final Report. California Youth Authority, April 1975.

Johnson, R. Culture and crisis in confinement. Lexington, Mass.: D. C. Heath & Co., 1975.

Johnson, R. Informal helping networds in prison: The shape of grass roots correctional intervention. Journal of Criminal Justice, 1979, 7, 53070.

Johnson, R., & Dorin, D. Dysfunctional ideology: The black revolutionary in prison. In Szabo, D., & Katzenelson, S. (Eds.) Offenders and corrections, New York: Praeger, 1978, 31-52.

Johnson, R., & Price, S. The correctional officer in today's prison: Human service, human community, and the human environment of prison. Paper presented at the 109th Congress, American Correctional Association, August 1979.

Johnson, R., & Price, S. The care and feeding of the correctional officer. Correctional Training, October, 1979, 3-5.

Joint Commission on Correctional Manpower and Training. Targets for inservice training. Washington, D. C.: Correctional Briefings #7, 1967.

Joint Commission on Correctional Manpower and Training. A time to act. Washington, D. C.: October, 1969.

Joint Commission on Correctional Manpower and Training. Perspectives on correctional manpower and training. Washington, D. C.: January, 1970.

Katrin, S. E. The effects on women inmates of facilitation training provided correctional officers. Criminal Justice and Behavior, 1974, 1, 5-12.

Kentucky Mental Health Manpower Commission. Community resource training for Kentucky probation and parole officers. Report of a Pilot Demonstration Project, 1974.

Kirk, R. E. Experimental design: procedures for the behavioral sciences. Belmont: Brooks/Cole, 1969.

Kirkpatrick, D. L. Evaluating training programs: Evidence vs. proof. Training and Development Journal, November, 1977, 9-12.

Law Enforcement Assistance Reform Act of 1979. S. 241, 96th Congress, 1st session May 21 (legislative day, April 9), 1979.

Lott, D. C. Evaluating to reduce training costs. Training and Development Journal, January, 1967, 38-41.

Mager, R. F. Goal analysis. Belmont: Fearon Publishers, 1972.

Mager, R. F. Preparing instructional objectives (2nd ed.). Belmont: Fearon Publishers, 1975.

Maloney, D. M., Phillips, E. L., Fixsen, D. L., & Wolf, M. M. Training techniques for staff in group homes for juvenile offenders: An analysis. Criminal Justice and Behavior, 1975, 2, 195-215.

Mandel, N. Making corrections a community agency. Crime and Delinquency, 1973, 17, 281-288.

Martinson, R. M. What works?: Questions and answers about prison reform. Public Interest, 1974, 35, 22-55.

Maryland Correctional Training Commission. Third Annual Report to the Governor, the Secretary of Public Safety and Correctional Services, and Members of the General Assembly, July 1973-July 1974.

Maryland Governor's Commission on Law Enforcement and Administration of Justice. Comprehensive plan for Maryland criminal justice training programs. State of Maryland Criminal Justice Report, June, 1976.

Mathieson, T. Across the boundary of organizations. California: Glendessary Press, 1972.

Matthews, C. V. Final report to the Connecticut Department of Corrections: An evaluation of the joint training academy. Carbondale: Southern Illinois University, Center for the Study of Crime, Delinquency, and Corrections, November, 1973.

Maxim, P. Treatment-custody staff conflicts in correctional institutions: A re-analysis. Canadian Journal of Criminology and Corrections, 1976, 18, 379-386.

McConkie, M. L. The role of interpersonal trust in correctional administration. University of Georgia, Corrections Division, Institute of Government, September, 1975.

McCormick, E. G. Job and task analysis. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

Milgram, S. Obedience to authority, New York: Harper, 1975.

Miller, W. B. Inter-institutional conflict as a major impediment to delinquency prevention. Human Organization, 1958, 17, 20-23.

Miller, W. B. Ideology and criminal justice policy: Some current issues. The Journal of Criminal Law and Criminology, 1973, 64, 141-162.

Moos, R. F. Correctional Institutions Environment Scale Manual. Palo Alto: Consulting Psychologists Press, Inc., 1974.

Morton, J. B. Southeastern Correctional Management Training Council. Final Report. (Discretionary Grant #74-ED-04-0009), December, 1975.

Myers, J. Y. Fundamentals of experimental design. Boston: Allyn and Bacon, 1972.

National Advisory Commission on Criminal Justice Standards and Goals. Corrections. Washington, D. C.: January, 1973.

National Advisory Network on Correctional Manpower, Training and Development. Centre for Social Welfare Studies, Carleton University. National Corrections Task Bank (Correctional Manpower Development Project). Ottowa, Ontario, 1976.

National Council on Crime and Delinquency. Probation and Parole Directory, United States and Canada (17th ed.). Hackensack, 1976.

National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice. The National Manpower Survey of the Criminal Justice System, Volume Three, Corrections. Washington, D.C.: September, 1978.

Nelson, E. K. Strategies for action in meeting correctional manpower and program needs. Crime and Delinquency, 1966, 12, 221-226.

Newstrom, J. W. Catch-22: The problems of incomplete evaluation of training. Training and Development Journal, November, 1978, 22-24.

Nosin, J. A. Pre-service and in-service training of Georgia correctional personnel: An interim evaluation. Atlanta: Planning and Evaluation Section, Research and Development Division, Department of Corrections/Offender Rehabilitation, January, 1975.

Nuehring, E. M. The character of interorganization task environments -- community mental health centers and their linkages. Administration and Society, 1978, 9, 425-442.

O'Leary, V., & Duffee, D. Correctional policy: A classification of goals designed for change. Crime and Delinquency, 1971, 17, 373-386.

Osborn, R. N., & Hunt, J. G. Environment and organizational effectiveness. Administrative Science Quarterly, 1974, 19, 231-245.

Paddock, A. L., & McMillin, J. D. Vienna staff training project. Final Report. Carbondale: Southern Illinois University, Center for the Study of Crime, Delinquency, and Corrections, June, 1972.

Parsons, T. Suggestions to a sociological approach to a theory of organizations. Administrative Science Quarterly, 1956, 1, 63-69.

CONTINUED

1 OF 2

Patton, M. Q. Utilization focused evaluation. Beverly Hills: Sage Publications, Inc., 1978.

Peretti, P. O., & Hooker, M. Social role self-perceptions of state prison guards. Criminal Justice and Behavior, 1976, 3, 187-196.

Quay, H. C., & Johnson, V. S. Training in correctional treatment techniques. Final Report. (Grant #73-Ed-99-0011). Coral Gables: University of Miami, Program in Applied Social Sciences, 1976.

Robinson, P. Delaware--Division of Adult Corrections--Staff development and in-service training--Project evaluation, August, 1975.

Rossi, P. H., Freeman, H. E., & Wright, S. R. Evaluation: A systematic approach. Beverly Hills: Sage, 1979.

Schon, D. Beyond the stable state. London: Temple Smith, 1971.

Scott, R. J., Cienek, R. P., & Evans, R. Correctional officer training in Virginia: A final report. (Report #2, Correctional Training Evaluation Project). Virginia Commonwealth University, Department of Administration of Justice and Public Safety, August, 1977.

Scriven, M. An introduction to meta-evaluation. In P. D. Taylor, & D. M. Conley (Eds.) Readings in curriculum evaluation. Dubuque: William C. Brown, 1972.

Scriven, M. Evaluation perspectives and procedures. In W. J. Popham (Ed.) Evaluation in education: Current applications. Berkeley: McCutchan, 1974.

Sharp, R. E. Effects of in-service counselor training on counseling ability of juvenile probation officers. (Doctoral dissertation, LeHigh University, 1974).

Shover, N. Experts and diagnosis in correctional agencies. Crime and Delinquency, 1974, 20, 347-358.

Shover, N. A sociology of American corrections. Homewood: The Dorsey Press, 1979.

Smith, R. R., Milan, M. A., Wood, L. F., & McKee, J. M. The correctional officer as a behavior technician. Criminal Justice and Behavior, 1976, 3, 345-360.

Solomon, P., & Gardiner, J. Political obstacles to change in criminal justice agencies: An interorganizational perspective. In Criminal Justice Monograph: The change process in criminal justice. National Institute of Law Enforcement and Criminal Justice, U. S. Department of Justice, June, 1973.

Sparks, C. P. Job analysis under the new uniform guidelines. Technical Report 79-6). Houston, Personnel Research, August, 1979.

Starbuck, W. H. Organizations and their environments. In M. D. Dunnette (Ed.) Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

Terryberry, S. The evaluation of organizational environments. Administrative Science Quarterly, 1968, 12, 590-614.

The Bellringer. A periodic review of criminal justice. Washington, D.C.: National Criminal Justice Association, 14, August, 1979.

Thomas, J. E., & Williams, T. A. Change and conflict in the evolution of prison systems: Old dilemmas, emergent problems, and future strategies. International Journal of Criminology and Penology, 1977, 5, 349-365.

Thompson, J. D. Organizations in action. New York: McGraw Hill, 1967.

Thompson, J. D., & McEwen, W. J. Organizational goals and environment. American Sociological Review, 1958, 23, 23-31.

Toch, H. Is a "correctional officer," by any other name, "screw?" Criminal Justice Review, 1978, 3, 19-35.

Toch, H. Living in prison. New York: The Free Press, 1977.

Tools for Trainers. Carbondale: Southern Illinois University, Center for Study of Crime Delinquency and Corrections, Correctional Staff Training Officer Institutes (OLEA Grant No. 241), 1978.

Treger, H. The reluctance of the social agency to work with the offender. Federal Probation, 1965, 29, 23-27.

Trist, E. A concept of organizational ecology. Paper presented to the three Melbourne Universities, July 1976.

U. S. Commission on Civil Rights. Civil and human rights in Oregon state prisons. Oregon Advisory Committee, 1976.

U. S. Department of Justice. Draft federal standards for corrections. Washington, D. C.: June, 1978.

U. S. Department of Labor. Handbook for analyzing jobs. Washington, D. C.: U. S. Government Printing Office, 1972.

Weiner, R. I. Some organizational and systemic constraints to change: A case study of the D. C. Department of Corrections: Washington, D. C.: Center for the Administration of Justice, The American University, 1973.

Weiner, R. I. Sociometric analyses of interorganizational relations: An exploratory study of factors which enhance or hinder cooperation between correctional and community service organizations. (Dissertation, University of Maryland, School of Social Work and Community Planning, 1977).

West Virginia-comprehensive training program for correctional personnel. Final Report. West Virginia University, Committee on Crime, Delinquency, and Law Enforcement. November, 1969-May, 1970.

White, M. F., & Dean, C. W. Problems in operationalizing theoretical ideas in correction. Journal of Research in Crime and Delinquency, 1969, 6, 87-98.

Witherspoon, A. D. A behavioral observation index designed to evaluate training of correctional officers in a prison setting. (Masters thesis, Auburn University, August, 1971).

END