If you have issues viewing or accessing this file contact us at NCJRS.gov.

FINAL TECHNICAL REPORT

NATIONAL EVALUATION PROGRAM
PHASE I ASSESSMENT OF

POLICE MANAGEMENT TRAINING PROGRAMS

84176

U.S. Department of Justice
National Institute of Justice

84176

This document has been reproduced exactly as received from the person or organization originating it. 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 National Institute of Justice.

Permission to reproduce this copyrighted material has been granted by
Public Domain/NIJ/LEAA
U.S. Department of Justice
to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permission of the copyright owner.

Prepared under contract number J-LEAA-021-78 awarded to Macro Systems, Inc., by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice. Points of view or opinions in this document are those of the authors and do not necessarily represent the official positions or policies of the U.S. Department of Justice.

FINAL TECHNICAL REPORT

NATIONAL EVALUATION PROGRAM
PHASE I ASSESSMENT OF
POLICE MANAGEMENT TRAINING PROGRAMS

Presented to:

National Institute of Justice
Law Enforcement Assistance Administration
U.S. Department of Justice

Prepared by:

James G. Ross
Mark T. Shanley
John L. Williams

Macro Systems, Inc.

October 24, 1980

TABLE OF CONTENTS

		Page Number
	ACKNOWLEDGMENTS	i
	SUMMARY	ii
ONE:	INTRODUCTION	1
	PART 1: PROGRAM DEVELOPMENT	11
TWO:	PROGRAM DEVELOPMENT PRACTICES	13
THREE:	FACTORS AFFECTING PROGRAM DEVELOPMENT	52
	PART 2: EVALUATION AND MANAGEMENT OPTIONS	76
FOUR:	PROGRAM MODELS	79
FIVE:	PROGRAM EVALUABILITY	117
SIX:	SINGLE PROGRAM EVALUATION	137
	PART 3: RESEARCH INDICATIONS AND GENERAL FINDINGS	183
SEVEN:	EVIDENCE OF PROGRAM EFFECTIVENESS	185
EIGHT:	RESEARCH APPROACHES TO NATIONAL POLICY QUESTIONS	194
NINE:	CONCLUSIONS AND RECOMMENDATIONS	203

TABLE OF CONTENTS (continued)

Appendices

1. DEFINITIONAL ISSUES AND EXPLICATION OF DEFINITION
2. METHODOLOGY
3. PROGRAMS VISITED
4. SITE VISIT REPORT OUTLINE
5. POLICE MANAGEMENT TRAINING PROGRAM SURVEY--NATIONAL EVALUATION PROGRAM PHASE I ASSESSMENT OF POLICE MANAGEMENT TRAINING
6. PROGRAM PERFORMANCE MEASURES
7. GUIDELINES FOR WRITING ACTION ITEMS
8. SELECTED BIBLIOGRAPHY

INDEX OF EXHIBITS

		Page Number
1.	EXPECTATIONS HELD FOR POLICE MANAGEMENT TRAINING PROGRAMS	5
2.	REPORT ORGANIZATION	8
3.	CORE STEPS IN THE GENERIC PROCESS OF TRAINING PROGRAM DEVELOPMENT	14
4.	SETTING TRAINING GOALS BASED ON BALANCED INPUTS	15
5.	CONDUCTING AN ASSESSMENT OF TRAINING NEEDS	18
6.	SETTING TRAINING OBJECTIVES	20
7.	TRANSLATING OBJECTIVES INTO SUCCESS CRITERIA	25
8.	DESIGNING A PROGRAM TO SERVE OBJECTIVES	26
9.	PRETESTING TRAINEES AND/OR THEIR DEPARTMENTS	34
10.	IMPLEMENTING THE PROGRAM DESIGN	36
11.	PERFORMING IN-PROGRAM EVALUATIONS	43
12.	PERFORMING FOLLOW-UP EVALUATIONS	45
13.	USING EVALUATION RESULTS	47
14.	ORIENTATION CHART TO 14 MODELS	81
15.	PRE-SERVICE AND INITIATORY COMPLIANCE MODEL	83
16.	REFRESHER AND UPDATE COMPLIANCE MODEL	86
17.	CORRECTIVE COMPLIANCE MODEL	88
18.	SYSTEMATIZED POLICING PRESCRIPTIVE MODEL	90
19.	STATE-OF-THE-ART PRESCRIPTIVE MODEL	92
20.	ADAPTATION PRESCRIPTIVE MODEL	95
21.	NON-EXPERIENTIAL PARTICIPATIVE MODEL	97

INDEX OF EXHIBITS (continued)

		Page Number
22.	EXPERIENTIAL PARTICIPATIVE MODEL	100
23.	"GREASING THE SKIDS" AUXILIARY MODEL	102
24.	CERTIFICATION AUXILIARY MODEL	105
25.	NETWORK AUXILIARY MODEL	107
26.	ANOINTING AUXILIARY MODEL	109
27.	DEPARTMENTAL DECISIONMAKING AUXILIARY MODEL	112
28.	CRITICAL MASS AUXILIARY MODEL	114
29.	IS THE PROGRAM WELL DEFINED?	119
30.	IS THE PROGRAM ACCEPTABLE?	123
31.	IS THE PROGRAM VALID?	124
32.	IS THE PROGRAM PLAUSIBLE?	127
33.	IS THE PROGRAM'S DATA SYSTEM FEASIBLE?	130
34.	IS THE PROGRAM'S DATA SYSTEM RELIABLE?	131
35.	ARE THE PROGRAM'S INTENDED USES OF DATA PLAUSIBLE?	132
36.	RANGE OF QUESTIONS FOR PROGRAM EVALUATION	138
37.	SAMPLE QUESTIONS THE ACTION PLAN APPROACH CAN ADDRESS	140
38.	BEHAVIORAL RECORD SHEET	154
39.	NONBEHAVIORAL RECORD SHEET	155
40.	INDIVIDUAL TRAINING MATRICES	157
41.	OUTCOMES OF TRAINEE ATTEMPTS AT IMPLEMENTING NEW BEHAVIROS (POLEX)	160
42.	PROBLEMS ENCOUNTERED BY TRAINEES IN IMPLEMENTING NEW BEHAVIORS (POLEX)	161

INDEX OF EXHIBITS (continued)

		Page Number
43.	NEW BEHAVIORS REPORTED BY TRAINEES (POLEX)	163
44.	EXTENT OF PROBLEMS ENCOUNTERED BY TRAINEES IN IMPLEMENTING NEW BEHAVIORS (POLEX)	164
45.	NATURAL GROUPINGS OF NEW REPORTED BEHAVIORS (POLEX)	166
46.	TRAINEE JUDGMENTS OF IMPACT FOR REPORTED BEHAVIORS (POLEX)	167
47.	KNOWLEDGE AND ATTITUDE CHANGE (POLEX)	169

ACKNOWLEDGMENTS

Over the course of this National Evaluation Program Phase I Assessment of Police Management Training Programs, we surveyed more than 200 programs by mail, visited 16 programs around the country, experimented with innovative evaluation approaches in three more, maintained contact with the state directors of law enforcement training in their POST agencies, sought the advice and counsel of experts in management training and training evaluation, and asked many people to critique draft reports.

Hundreds of people offered us valuable information and sidelights. They gave us literally thousands of hours of time. A study of this complexity and national scope could not have been completed without their help. We wish to express our appreciation to all who cooperated with our study and lent us their special expertise. Specifically, we thank:

- The state directors of law enforcement training, who explained their agencies' functions, gave us entree to state-certified programs, and kept us posted on state-level developments
- The directors of the 16 programs that we visited (listed in Appendix 3), who opened their doors to us and took time to explain their programs' development, goals, and activities
- The directors of the over 200 programs that completed a mail survey on their program development practices
- The directors of the three programs that let us test innovative evaluation approaches on them, for their continuous support and collegiality over an extended time period
- The instructors and trainees from all the programs we worked with, for their candor, receptiveness, interest, and incisive remarks
- LEAA staff, including Dale Beerbower, Larry Giddings, Jan Hulla, Lou Mayo, and above all our monitor, Phil Travers, for their critical insight, confidence, and guidance
- U.S. Office of Personnel Management (OPM) staff, including Alan Breitler, Garland Phillips, and Ruth Salinger, who explained their training evaluation approaches and advised our own evaluation efforts with a spirit of collegiality
- All who reviewed and critiqued our draft reports, including the LEAA and OPM personnel named above, and many others, such as Bob Brenner, Pat Gallagher, Esther Heffernan, Jim Horner, Joel Pate, Barbara Price, Geri Schaeffer, John Velier, and George Williams.

We hope that those who contributed to the study view the final products as a fair return on their investment.

SUMMARY

The National Institute of Justice (formerly part of LEAA) looked on police management training programs as vital enough to fund a National Evaluation Program Phase I Assessment about them. There were strong reasons for wanting better information about the development, management, and evaluation of these programs. The number run by state and local agencies has risen rapidly in recent years without benefit of the experiences of others. The question of management training evaluation is controversial, with some people arguing that gut feel is a sufficient barometer of program success and others trying to make programs act more accountably by demonstrating their payoff and worth more concretely. This was also an important topic because police management training programs play an indispensible role in the development of police managers and in the introduction of new concepts and practices into policing.

This study explored a number of questions, several of which we discuss briefly below.

1. <u>How closely do police management training programs follow the industrial model of training program development</u>?

   Our first reaction was to say that programs are developed "by the seat of the pants," hardly in a deliberate and systematic fashion. For example, program developers and operators set goals largely without substantial input from user groups; do little or no formal needs assessment; typically skip over the identification of performance deficiencies and often pass off topical interest surveys as need assessments; do not set consistently clear objectives; do not identify criteria that indicate, in measurable terms, the areas in which change is desired and the intended extent of change; and so forth. Our second reaction, however, took the broader view that practices in development of police management training are not unreasoned, that the process is deliberate, phased, and rather systematic but corresponds inconsistently from point to point with the industrial model.

2. <u>How systematic can program development reasonably hope to be</u>?

   There are factors external to programs and largely beyond the program developer's control that limit how closely developmental practices can correspond to any chosen system. At each major developmental juncture, five factors affect program development: funding, legal requirements, organizational environment, community environment, and the ready availability of resources and materials from prior programs. Other factors affect development at isolated points. It seems as if the program developer's options are ringed by a multitude of forces

beyond his control. Most program managers and operators have rightly concluded, therefore, that a high level of adherence to the industrial model is infeasible now and is unlikely to become more feasible in the near future. We add that most program managers and operators still have the capacity to make isolated interim changes that will make their programs more manageable. We also point out that a program developer's attitudes toward the value of systematic development can play a key role in activating and maintaining certain obstacles to more systematic practice.

3. Do police management training programs follow a single model or several different ones?

Police management training programs take many forms, ranging in the functions they seek to serve and in the means used to achieve chosen ends. Of the 14 police management training models that we identified, no single model was either fully articulated or unequivocally espoused by the programs we observed. Much of the model mixing we found was officially recognized, set forth in public descriptions of programs, fit together comfortably, and was quite legitimate. A lot of mixing stemmed from a lack of coordination that resulted in a "smorgasbord" type of program that pointed trainees in no clear direction. Much of it stemmed inevitably from the different responsibilities and stakes that people have in a program. But this mix or "coexistence" among several models in a single program often produced ambiguity about the model or models in which the programs operated. As a result, people with different functional responsibilities and stakes developed divergent notions of how the program should operate and did not act in concert. The variation among models and the phenomenon of model mixing have major implications for how programs should be managed and evaluated.

4. How do programs stack up against the criteria for operation of an evaluable program, i.e., one in which resources can be effectively managed and an evaluation conducted with reasonable chance of being useful?

Programs generally fall far short of the evaluability criteria, even when an appropriate evaluation technology happens to be available. For example, program expectations and activities are not at all well defined. Significant gaps and contradictions exist between program descriptions offered by policy-makers, program managers, and program operators. Program expectations are often implausible, in light of general educational and training theory, the extent and types of resources brought to the program, the manner in which resources are used, and evidence of program relevance and effectiveness from prior program experience. Based on such conditions, it seems that most police management training programs are far from optimally managed and that a major investment in evaluation is not what they need.

5. What should a program look at in an evaluation, when one happens to be appropriate?

The key question in program evaluation is not whether trainees liked a course or learned anything new from it but whether training made any difference in later job behavior. Few programs try to answer this question. We examined three approaches to it that have wide applicability and require relatively low investment. The Action Plan Follow-Up Approach asks whether trainees carry out on their jobs the intentions they took away from training. The Proficiency Analysis Approach asks whether training improves job proficiency in ways that the organization can use. Simplified Cost-Benefit Analysis asks whether training is a reasonable investment to make, based on its return to the organization in increased job proficiency. All three approaches have shown their usefulness in industry and government, but none has been extensively used to evaluate police management training.

6. To what extent are police management training programs effective?

The only evidence most programs can point to, reaction surveys and course final examinations, let us reach only two conclusions. First, trainees leave nearly all programs, regardless of variations in quality among them, satisfied that their time has not been wasted. Second, trainees leave programs with as much as or more knowledge of rudimentary concepts dealt with during a course than they carried into it. The evaluation literature offers scant evidence that the knowledge and skills trainees learn are relevant to their personal needs, will be usable in their agencies, and have measurable payoff in greater individual proficiency and organizational effectiveness. The evidence seems contradictory at points and allows us to draw few conclusions about overall effectiveness.

7. What are the important national policy questions in the area?

There are roughly a dozen questions on which the future structure, practices, and orientation of police management training hinge. The key issues include:

- The relative effectiveness and cost-effectiveness of residential and nonresidential programs
- The effects of a network of program graduates on the management practice, career development, and interdepartmental relations of network members
- The relative influence of experiential and nonexperiential training methods on later job behavior
- The effects of attendance at major national programs on the rapidity of career development and on the height of terminal positions occupied

- The relative influence of credentials and greater proficiency on the career patterns of major national program graduates
- The effects of pretraining personalized learning objectives developed cooperatively by trainee and supervisor on later implementation levels
- The effects of departmental pre-service and refresher courses on later compliance with departmental policies and regulations
- The relative receptiveness of small and large departments to new practices
- The effects of management training as a tool in departmental decision-making on the smoothness of organizational change
- The effects of management training in MBO on department-wide changes in management philosophy
- The inevitability of organizational change once a critical mass of similarly trained and attuned managers has been attained

Based on these conclusions, the study also directed recommendations for program management, operation, and evaluation to three groups: user agencies, program operators, and program managers.

One: Introduction

A comprehensive study of police management training programs is important and timely for several reasons. First, never before have so many police managers been expected to display sophisticated management skills and know-how. These programs, if effective, can play a major role in upgrading administrative and management capabilities of police managers faced with escalating job demands and system pressures. Second, because the police organizational structure tends to be para-military and authoritarian, it resists the introduction of innovations made by other police agencies and of modern management techniques and principles. Formal training programs are one of the few accepted, natural points of entry for the introduction of knowledge that could change the orientation and organizational structure of police management. Third, to meet training needs, state and local agencies have increasingly tried to start their own programs in recent years. This multiplication of programs has occurred largely without benefit of information about how programs are developed and how well they work elsewhere in the country. Fourth, there is much controversy over the question, "can management training comply with rigorous standards of evaluability?" This makes it opportune to test the position that management training is essentially unevaluable and that "gut feel" is a sufficient barometer of program success against the increasing demand that management training programs demonstrate more concretely their payoff and worth.

In this introduction, we first elaborate on the need to study police management training programs and lay out the broad range of expectations that people hold for these programs. Then we set forth the study's purposes, explain the differences between our approach and traditional evaluations, describe the audience to whom this report is directed, and list our data sources. Finally, we outline the report's contents, the questions it seeks to answer, and the ways its information can be used.

Throughout this report, we define a police management training program as:

AN INSTRUCTIONAL SEQUENCE
OFFERED OUTSIDE A DEGREED ACADEMIC PROGRAM
FOR THE PURPOSES OF UPGRADING ONE OR MORE ASPECTS OF
    SUPERVISORY OR MANAGEMENT PERFORMANCE
FOR THE ULTIMATE BENEFIT OF A POLICE AGENCY
TO CURRENTLY ACTIVE OR SOON TO BE COMMISSIONED POLICE
    SUPERVISORS OR MANAGERS
WHO OPERATE ON THE STATE OR LOCAL LEVELS.

A detailed version of this definition and an analysis of definitional issues are in Technical Appendix 1.

A. WHY THE NEED TO STUDY POLICE MANAGEMENT TRAINING PROGRAMS?

The importance and timeliness of this study rest on the four points we made above, about upgrading the police manager's capabilities, changing the police organizational structure, identifying state and local program alternatives, and identifying program evaluation options. Briefly, we now place each in perspective.

1. Upgrading the Police Manager's Capabilities. For at least fifty years, one national study after another has decried the typical police manager's lack of administrative and managerial skills. The Wickersham Report in 1931, the Police Executive Report in 1975, and the National Manpower Survey in 1978 all made similar observations in this regard. Over this time, the demands placed on the police manager have seemed to increase geometrically. Today's police manager is under fire to conduct a disciplined system of objectives and priorities and to expand departmental capabilities for audit and performance measurement. The manager is swept up in the trend toward reassessing organizational objectives and a growing need to analyze how his personnel system should best be structured and managed. He needs the knowledge, skills, and attitudes for effectively responding to these and other pressures. It is doubtful that these skill deficits can be met just through informal working contacts with others who have also fish-laddered their way through the ranks and are, by their own admission, still little more than promoted policemen. Management training is probably the most efficient way to upgrade individual managerial skills.

2. Changing the Police Organizational Structure. Police agencies are known for their resistance to change. Some people say this is because police agencies tend to be para-military and authoritarian. They are called para-military because they often rely on a formal and rigid chain-of-command, detailed rules and procedures, the insulation of roles, and bureaucratic techniques. They are called authoritarian because they involve the use of threats related to physical/economic needs, decision-making at the apex of the organization with little rank-and-file participation, exclusively downward communications, excessive competition among peers and between subordinates and superiors, and opposition to the organizational goals by a strong informal organization.

Whatever the reasons for this resistance, police agencies do not take to change easily. This obstructs the introduction of innovations made by other police agencies and the general adaptation of business management principles to the police environment. Management training is one of the few accepted, "natural points of entry" for introducing new concepts to police agencies. An authority named Yin argued that the best way to introduce new knowledge is through "organizational linkage already embodied in the complexity of the practitioner's world."1/ The forum of a management training course is, therefore, one of the few accepted means for influencing the organizational structure and orientation of police agencies.

1/ See R.K. Yin, R & D Utilization by Local Services (Santa Monica, California: Rand Corporation, 1976), pp. viii-ix.

3. Identifying State and Local Program Alternatives. There has been much recent experimentation on the state and local levels with police management training without much information about the experiences of others. This multiplication of programs seems to have occurred for several reasons. More and more people recognize the managerial shortcomings of police managers. Beliefs about the desirable type of police manager have become more varied, requiring additional programs to reflect these beliefs. State and local authorities have demanded programs better geared to state and local needs and concerns. The state-level police standards and training agencies (POSTs) have broadened their influence and sought to strengthen their offerings. LEAA has made funds available for program development through SPAs and POSTs. State and local programs also cost less to operate than do out-of-state, residential programs. The lower per-trainee cost permits more officers to be exposed to training. Officers attending training nearby can be called back in emergencies or to resolve coverage problems. If officers can commute to training, there is less strain on family life. Whatever the reasons for this expansion of state and local training options, much of this has occurred without benefit of knowledge about how programs are developed, managed, and evaluated elsewhere in the country or about what practices ought to be emulated.

4. Identifying Program Evaluation Options. There is controversy brewing over whether management training in any field is really evaluable. One camp argues that it is impossible or unimportant to articulate the objectives of management training clearly and to evaluate its effects systematically. It sees management training as "something that will take care of itself" or as "an act of faith." It regards "gut feel," or trainee satisfaction, as the only feasible measure of program success. To some degree, this attitude validly recognizes that management training is not as amenable to evaluation as a basic training course in defensive driving because it is harder to define what "success" means and harder to find measures of success. The other camp, however, subjects management training to greater scrutiny and insists that it be more accountable by showing its payoff and worth to the organization more concretely. Mindful of this controversy and the continuing spread of state and local programs, it is timely to explore how and when programs can be usefully evaluated.

B. WHAT ARE POLICE MANAGEMENT TRAINING PROGRAMS EXPECTED TO ACHIEVE?

Police management training programs are expected to serve as means to a wide variety of ends. This variation reflects different concepts of the police manager's knowledge and skill requirements, of the constraints that the police manager must know how to deal with, and of how careers can be shaped and law enforcement practices changed through training. Programs have been expected to result in:

- Performance of the police manager's duties in closer compliance with the responsibilities and policies of the particular police agency
- Improved individual and agency confidence and morale
- Greater individual and agency effectiveness by adaptation of business management practices in the systematization of police administration

- Movement of police managers away from authoritarian and toward more participative management practices, shown in more open communications, less reliance on sanctions as a motivator, and greater rank-and-file input into decision-making

- Maintenance of a mutually reliant network of program graduates, shown in continuing contact and cooperative problem-solving among police managers in diverse agencies

The range of expectations held for police management training programs is much wider than this, however. Exhibit 1 shows more fully the range of outcomes and impacts people want to see resulting from police management training. Programs are not and need not be uniform; variation in expectations reflects the difference among the activities in which trainees between one program and another take part.

C. WHAT ARE THE STUDY'S PURPOSES?

When this study was first planned, it had two broad purposes. The first purpose was to describe the evaluation options open to police management training programs. This would involve constructing models of police management training, identifying points at which evaluation measurements could be taken, portraying the obstacles that impede useful evaluation and effective program management alike, and outlining evaluation designs that require relatively low investment. The second purpose was to develop a research agenda on the subject of police management training. This would involve an assessment of available evidence about program effectiveness and the development of research designs to address those issues for which interest is high but useful information low.

To these two broad purposes a third was added, upon urging by LEAA staff and POST directors. This was to assess the relationship between industry's structured and deliberate process for training program development and that actually followed by police management training programs. This would involve outlining the prescribed developmental steps, comparing them with current practice in development of police management training programs, and then analyzing the factors that facilitate or impede systematic program development. The three purposes, although interrelated, were distinct. In effect, we were almost doing three separate studies.

The study's ultimate purpose was not so much to influence the quantity and quality of program evaluations as to offer structured ways to think about what a given program is designed to accomplish and to identify ways of adjusting a program that will make it more manageable. Some early reviewers have referred to this report as a "frame of reference for future planning" and a "manual for rational program development."

D. HOW DOES THIS STUDY'S APPROACH DIFFER FROM TRADITIONAL EVALUATIONS?

In traditional evaluations, a design is often planned and implemented in isolation from decision-maker interests and operating program realities. At often great expense, they produce reams of data that are of little or no use to

EXHIBIT 1

NEP/Police Management Training

EXPECTATIONS HELD FOR POLICE MANAGEMENT TRAINING PROGRAMS

Performance of the police manager's duties in closer compliance with the responsibilities and policies of the particular police agency

Maintenance of executive control over line managers in particular police agencies

Improved individual and agency confidence and morale

More thoughtful self-examination by police managers of their current and alternative managerial roles

Greater individual and agency effectiveness by adaptation of business management practices to the systematization of police administration

Greater receptivity to police management research and experimentation with innovative and demonstrably effective practices tried by other police agencies

Greater organizational capability to anticipate and adapt to changes in Federal and state laws and regulations

General expansion of the range of options normally considered by a police manager, due to heightened awareness of the need for responding to immediate situational factors

Movement of police managers and agencies away from authoritarian and toward more participative management practices, shown in more open communications, less reliance on sanctions as a motivator, and greater rank-and-file input into decision-making.

Improved individual capability in specific functional areas, such as budget preparation and personnel management

Development of attitudes and personal skills more in line with changes in police management systems

Greater police professionalism, shown in increased enrollment in advanced education and training

Reduced managerial turnover paralleled by the attraction of more qualified individuals to managerial jobs

Accelerated career development and achievement of departmental status in recognition of highly prized credentials obtained in training

Provision of a conduit for critical feedback to executives in agency decision making

Maintenance of a mutually-reliant network of program graduates, shown in cooperative problem-solving among police managers in diverse agencies

Development of a "critical mass" of trained police managers capable of initiating and guiding large-scale organizational change

Improved agency productivity and capacity to function in concert with other government and private sector organizations

program managers and operators. One of this study's reviewers compared traditional evaluations of police management training to "putting the cart before the horse." This reviewer added that, "Once we in law enforcement get our act together as to what management training should consist of, then the problem of evaluation should be addressed."

The main concern of this study, however, is helping to create conditions that permit both useful evaluation and improved program management. Our focus is upon helping program managers and operators to "get [their] act together as to what management training should consist of," how training resources can best be used, and when and whether training's results can usefully be measured. This practical approach rests on a simple assumption:

> When a program does not meet conditions for evaluability, its managers and operators will encounter difficulties in making it work, demonstrating how well it works, or both. Periodically answering certain preliminary questions not only leads to more useful evaluation designs but also helps identify roadblocks to effective program management and the documentation of a program's success.

The core of the evaluation approach we follow and endorse involves these steps:

- Determine how a program rates on the seven basic criteria for an evaluable program. For example, does program management define with reasonable completeness what is expected to happen in and result from the program? Is it plausible that the program will accomplish its purposes? Does the program actually in place validly represent program management's expectations? (See Technical Appendix 2 and Chapter Five for a full discussion of the evaluability criteria.)
- Identify the precise practices or conditions that obstruct or disable efforts to operate an evaluable program.
- Determine how the program could be adjusted to make it more evaluable (and, hence, more manageable).
- Proceed to collect evaluation data only to the extent that prior steps have established that evaluation will be a useful expenditure.

The person assigned evaluation responsibilities--who may be a professional evaluator but more likely works on the training staff or for an oversight/funding agency--can have several different responsibilities in this evaluation process, depending on the situation. The first is to document the roadblocks to evaluability--the program's shortcomings--as part of a detailed description of the program. The second is to work with program managers and operators to identify ways to remove roadblocks to evaluability by adjusting expectations, altering activities, or upgrading the information system. The third is to help program managers and operators to select a strategy for improving the program. The fourth is to help monitor the implementation of the strategy. The fifth is to design and implement an evaluation approach that meets the program's needs. With

the possible exception of the last evaluator responsibility, these evaluation activities do not require the evaluator to have special training or skills in traditional evaluation approaches.

E. TO WHOM IS THIS REPORT DIRECTED?

In conducting this study, seven sets of individuals expressed interest in our work. They are: evaluators; researchers engaged in scholarly investigation; academicians and others with an interest in training or management theory; professional organizations, including the American Society of Criminology, the International Association of Chiefs of Police, the National Organization of Black Law Enforcement Executives, the Police Executive Research Forum, and the Police Foundation; program managers (working in funding and/or oversight agencies); program operators; and decision-makers in police departments.

To meet varied interests, we have written two main reports: a Technical Report and a Summary Report. This Technical Report we have directed primarily to the first three sets of individuals--evaluators, researchers, academicians, and others with an interest in theory--and to persons in the other sets who have occasional need for comprehensive, detailed information on particular topics. To meet the needs of practitioners, the Summary Report condenses and refocuses those parts of the Technical Report that should be of most use in answering questions about oversight and operation of programs. Because much material in the reports has general relevance to those outside of law enforcement, we hope the reports will also find their way to other fields.

F. WHAT ARE THIS REPORT'S DATA SOURCES?

It is not essential to grasp the study's methodology fully to appreciate the substance of the report. For those with methodological interests, the methodology is described in Technical Appendix 2. Still, it is useful in reviewing any document to know the authors' sources. This report draws upon six major data sources:

- Ongoing consultation with experts in the field of police management training and management training evaluation
- Ongoing review of documents that deal with police management training or management training evaluation, including a limited number of program evaluations
- Preliminary telephone surveys of police standards and training agencies on the state level, SPAs, and selected programs
- Site visits to sixteen police management training programs, where we observed training, reviewed files, and interviewed program directors, trainers, other staff, and the trainees themselves
- A large-scale mail survey directed to 250 police management training programs nationwide (to which 90 percent of active programs responded)
- Experimental use at three programs of two evaluation approaches described in Chapter Six

EXHIBIT 2(1)

NEP/Police Management Training

REPORT ORGANIZATION

Chapter or Appendix Number	Chapter or Appendix Title	Data Sources (In Declining Order of Importance)	Principal Questions Asked	Potential Uses of Information
Chapter One	Introduction	. Other chapters of the report, or other study reports	. Why study police management training? . What are the study's scope and purpose? . How is the report organized?	. Determine whether the study is of interest. . Determine which reports are of greater interest. . Determine the priority with which specific parts should be read.
PART ONE				
Chapter Two	Program Development	. Site Visits . Mail Survey . Preliminary Telephone Surveys . Literature Review	. What generic steps does industry recommend that training programs follow in their development? . What are the rationales for each step? . Do police management training programs follow them?	. Assess how closely police management training programs in general follow industrial standards. . Assess how closely a particular program follows industrial standards. . Identify new approaches to program development.
Chapter Three	Factors Affecting Program Development	. Site Visits . Mail Survey . Preliminary Telephone Surveys	. What forces beyond a program's control can affect a program developer's ability to follow the developmental practices of his choice? . Are these forces so strong as to preclude following a reasoned developmental process?	. Assess the strength of forces that affect development of police management training programs generally. . Identify the forces that affect a specific program's developmental practices. . Assess the feasibility of overcoming the forces that affect program development.
PART TWO				
Chapter Four	Program Models	. Site Visits . Expert Consultation . Preliminary Telephone Surveys . Review of Program Brochures	. What different models do programs follow, distinguishing among them by their assumptions, activities, and objectives? . Do the models focus just on substance or also on how training is meant to impact upon agencies? . To what extent do programs clearly articulate a single model that they consistently follow?	. Identify the models a given program follows. . Identify the activities a program ought to consider to be consistent with its objectives. . Clarify how a course's substantive influence is meant to impact on the environment. . Identify ways that a program may need to clarify its model, to be more internally consistent.
Chapter Five	Program Evaluability	. Site Visits . Review of Evaluation Literature . Preliminary Telephone Surveys . Mail Survey	. What are the criteria for program evaluability? . What roadblocks impede program evaluability? . How frequently is each roadblock confronted? . What can be done to overcome these roadblocks and make programs more evaluable?	. Assess the relative potency of roadblocks to evaluability in programs generally. . Identify the roadblocks to evaluability that affect a particular program. . Develop a strategy for making programs in general more evaluable. . Develop approaches for making a particular program more evaluable.
Chapter Six	Single Program Evaluation Approaches	. Review of Evaluation Literature . Expert Consultation . Experimentation with the Approaches	. What are the evaluation questions that the typical program might want to answer? . What needs to be done before an evaluation is started? . How can evaluation show whether training makes any difference in later job behavior?	. Identify the evaluation questions germane to a particular program. . Determine whether to evaluate a program. . Implement a program evaluation focused on later job behavior.

EXHIBIT 2(2)

Chapter or Appendix Number	Chapter or Appendix Title	Data Sources (In Declining Order of Importance)	Principal Questions Asked	Potential Uses of Information
PART THREE				
Chapter Seven	Evidence of Program Effectiveness	. Review of Evaluation Literature . Site Visits	. In what ways have police management training programs been evaluated? . What conclusions about program effectiveness can be drawn from these evaluations? . What policy questions should be the focus of future research?	. Assess the strength of evidence about the effectiveness of programs in general. . Identify research questions that should be the focus of future research. . Build future research efforts on the informational gaps found in prior research.
Chapter Eight	Research Approaches to National Policy Questions	. Review of Research Literature . Review of Public Policy . Site Visits . Expert Consultation	. What types of policy implications could future research pose for public policy? . What types of research designs can be implemented to address specific policy questions?	. Determine the usefulness of further research on a particular question. . Implement research on a particular policy question.
Chapter Nine	Conclusions and Recommendations	. Entire Report	. In brief, what conclusions can be drawn from this study? . What actions should be taken on the basis of this study's findings?	. Reassess a particular program from several analytic vantage points. . Identify a broad range of actions that might be taken to improve police management training.
TECHNICAL APPENDICES				
Appendix One	Definitional Issues and Explanation of Definition	. Review of Training Literature . Expert Consultation	. What issues have to be resolved to define the topic? . In detail, what programs does the definition include/exclude?	. Appreciate the subtleties behind the definition. . Test particular programs to see if they fit the definition.
Appendix Two	Methodology	. Review of NEP Literature . Project Records	. What assumptions and tasks does the NEP require? . In detail, what data collection procedures did the study use? . What methodological problems did the study encounter?	. Appreciate the NEP requirements placed on the study. . Appreciate how an iterative process works. . Replicate and build on the study's methods. . Avoid any methodological problems that the study faced.
Appendix Three	List of Sites Visited	. Project Records	. Which programs did the study visit?	. Understand the range of programs visited.
Appendix Four	Site Visit Report Outline	. Project Records	. How was information collected on site visits organized?	. Appreciate how site visit data were used in the study. . Identify possible ways to organize data in future studies.
Appendix Five	Large-Scale Mail Survey	. Project Records	. What questions did the large-scale mail survey ask?	. Build future surveys on findings from and gaps in the NEP survey.
Appendix Six	Performance Measures	. Review of Training Evaluation Literature . Site Visits	. What types of measurements can be taken to evaluate program success?	. Appreciate that several different measurements for the same event are often possible. . Identify the evaluation measures for a particular program.
Appendix Seven	Action Plan Evaluation Forms	. Review of Training Evaluation Literature	. What forms and handouts should be used in the Action Plan Follow-Up Evaluation?	. Implement the Action Plan Follow-Up Evaluation.
Appendix Eight	Selected Bibliography	. Project Library	. What resources should be consulted for further information?	. Obtain and read the most relevant works on several topics.

The study's methodology was both phased and iterative. This means that we did not simply draw a set of hypotheses at the study's start, design a data collection strategy, collect the required data, analyze the data, and then stop. On the contrary, several means of collecting data were often used to develop progressively better information on a given question. Each task developed information that was used to sharpen the questions asked and to refine the products of other tasks. The advantage of employing a methodology that encourages "successive correction" like this is that one can explore terrain that was not envisaged at the study's onset while avoiding areas that turn out to be non-productive and unimportant.

G. HOW IS THIS REPORT ORGANIZED?

This report is organized in three parts, each of them centered around one of the study's purposes.

- Part One centers around the industrial model of program development. It contains two chapters, Two and Three. Two compares current practice with the industrial model of program development. Three examines the factors that affect whether program development follows a reasoned path.

- Part Two offers new ways to think about program evaluation (and, hence, program management). It contains three chapters, Four, Five, and Six. Four outlines the 14 models of police management training observed in the field. Five discusses the evaluability of police management training programs nationwide and explains a self-assessment process for improving a program's evaluability. Six describes three results-oriented evaluation approaches that require low investment but promise high informational yield.

- Part Three centers around completed evaluations and research. It contains three chapters, Seven, Eight, and Nine. Seven assesses what conclusions can be drawn from the available evidence about the effectiveness of programs. Eight sets forth ways of researching unanswered questions that have major policy implications. Nine summarizes conclusions and recommendations for the preceding eight chapters.

Exhibit 2 shows how the report is organized. We hope that this exhibit will help in deciding how best to use the report. It presumes that many readers will find selective reading to be more appropriate in meeting particular interests than straightforward cover-to-cover reading. To aid this selective reading, the exhibit includes each chapter and technical appendix by number and name, indicates the data sources used to write the chapter (in decreasing order of importance), lists the questions the chapter addresses, and notes different ways in which the chapter can be used. We have declined to indicate a "most appropriate user" chapter-by-chapter because most chapters may be used in several ways, by different users. Each reader can determine what uses happen to fall within the scope of his or her interests and influence.

PART 1: PROGRAM DEVELOPMENT

Industry's celebrated design and use of a generic and highly adaptable process for training program development has caused many people to ask, "How systematic can our own training programs hope to be?" The managers and operators of police management training programs are no exception to the rule. Concern has arisen in some quarters over whether these programs exhibit reasonable correspondence in their development with the industrial model. Deviation from this process is presumed by many to mean that programs are not evaluable and, hence, not manageable. In other quarters, it has been questioned whether it is realistic to demand that police training programs imitate the industrial model, especially in a "soft" area like management training.

This part of the report centers around the industrial model of program development by asking two questions:

1. What is the relationship between the industrial standards and actual practice in development of police management training? Step by step, how closely do these programs follow the industrial model?

2. What external factors influence how systematic program development practices can reasonably hope to be?

Chapter Two explores Question 1 by outlining ten core developmental steps and their rationales, breaking each core step down into two or more substeps with corresponding rationales, and then describing current practice in the context of the 42 resultant substeps. Centering this discussion around certain prescribed steps does not, however, constitute an affirmation of their validity. That is a larger issue, beyond our scope but strongly related to feasibility, which Chapter Three partially addresses. It is also beyond our scope to ask if industry closely follows the standards it has designed. With these understandings, we treat departures from the steps as empirical facts rather than as evidence of "mediocrity or regression," as one POST director said such comparisons seem to imply. We offer no judgment as to the relative effectiveness of procedures. We merely describe, from point to point, the consistencies and inconsistencies between the standards and current practice.

Chapter Three examines the factors external to programs and largely beyond the program developer's control that affect how closely program development practices can correspond to any chosen system. We look at certain factors that can affect program development at any point, from initial goal identification and needs assessment through evaluation of results. These pervasive factors include funding, legal requirements, organizational environment, community environment, and the ready availability of materials and resources from prior programs. Other factors affect program development at specific junctures. The homogeneity and

stability of the target population, for example, can influence the feasibility and usefulness of a needs assessment. Departmental coverage requirements and the priority that served agencies give to training can dramatically affect availability of the type of trainee for whom a course was designed. Chapter Three describes in detail how factors like these can affect how deliberate and systematic program development can reasonably expect to be.

Both chapters in Part 1 draw on four data sources. The overall developmental framework builds on carefully chosen works in training program development and evaluation, including works specific to police training. The descriptions of developmental practices and the factors affecting them draw on three sources: the semi-structured telephone surveys of POSTs, SPAs, and selected programs, conducted during the study's preliminary stages; intensive on-site observation of sixteen programs, review of program files, and interviews with their staffs, instructors, and trainees; and the national mail survey of nearly all active programs. The preliminary surveys gave sensitivity to issues, site visits an in-depth appraisal of developmental practices and constraints, and the national mail survey the grounds for generalizing on what the site visits showed. Despite inadequacies in how programs document their development and despite some inherent limitations on our own data sources, our findings are generally conclusive.2/ Current patterns in program development, and the net of factors that affect how systematic it can be, are both clearly spelled out.

2/ There was a marked tendency for respondents to the national mail survey to understate practices or conditions that might be seen as undesirable and to exaggerate those likely to be seen as desirable. We come to this conclusion based on cross-checks within the survey, comparison with site-visit observations, and consideration of other available information. The national survey data are still important as a relative indicator of practices, even if not as an absolute one.

Two: Program Development Practices

To consider the question, "Do police management training programs follow the industrial model of program development?," we first need a program development framework. No universally accepted set of procedures for training program development has been promulgated, but numerous guidelines suggest aspects of what a systematic process ought to involve. Each set of standards has its own emphases and details certain parts of the process while playing down others. One set of standards, for example, focuses on the steps for linking program development with the parent organization's general planning process to ensure organizational input and eventual "buy in." Others detail needs assessment procedures, design factors, or evaluation techniques. In Exhibit 3, we draw on several sources to display a reasonably balanced progression of core developmental steps and related rationales.3/

We do not mean either the core steps or their substeps to be viewed as definitive. They merely provide a framework for ordering a description and analysis of practice. Each reader should have changes to recommend based on personal interpretation of the industrial model. The steps can be adapted to accommodate the range of auspices and inter-organizational arrangements under which training occurs. They can be collapsed or combined when immediate need and resources so require.

As presented here, the steps move with deliberation from a broad determination of goals, to assessment of needs for the general target population, to identification of in-program objectives under the program's control, to operationalization of these objectives, and only then to construction of a curriculum reflecting objectives, to pretesting of trainees, and then to flexible implementation of the program design so that actual needs of the immediate training class are addressed, and on to evaluation. In practice, conditions may not require or permit following the steps closely; and, in some instances, practices may shift over time. There can be vast differences between the original, trial-and-error developmental process and the later, more formalized process as programs move into higher gear.

3/ The steps were synthesized from: W.S. Bollinger and K.O. Vezner, Police Training Evaluation: A Systemic Approach (Toledo: Toledo/Lucas County Criminal Justice Supervisory Council, 1975), pp. 1-12; I.I. Goldstein, Training: Program Development and Evaluation (Monterey, California: Brooks/Cole Publishing Company, 1976); D.L. Kirkpatrick (ed.), Evaluating Training Programs (Madison, Wisconsin: American Society for Training and Development, 1975); V. Strecher (ed.), "Police Personnel and Training," in Cyclopedia of Policing: A Feasibility Study, Volume 13 (Los Angeles: Loyola University of Los Angeles, 1973), pp. 1364-1366; and A Process for the Evaluation of Training (Washington, D.C.: U.S. Civil Service Commission, Training Leadership Division, April 1978).

EXHIBIT 3

NEP/Police Management Training

CORE STEPS IN THE GENERIC PROCESS OF TRAINING PROGRAM DEVELOPMENT

	Step	Rationale
Step 1	Set broad training goals based on balanced inputs from those to be affected by a proposed program.	Training goals express the general intent behind and direction for program activities, and suggest the scope of needs to be further assessed. When based on inputs from all parties to be affected by a program, they ensure that training reflects those conditions and demands that the trainee will face on the job and that program standards are coordinated with the operational standards of other criminal justice and community agencies.
Step 2	Conduct a needs assessment to refine the dimensions where discrepancies exist between current and desired levels of individual performance and/or organizational functioning.	Achievement of program goals requires determining with precision what performance deficiencies exist, specifying those changes in individual and/or organizational performance that are consistent with goals.
Step 3	Develop training objectives to meet identified needs in the context of overall program goals.	Objectives express the training outcomes seen as indicating satisfactory attainment of desired individual performance and/or organizational functioning in the context of overall program goals.
Step 4	Translate objectives into measurable success criteria reflecting, at a minimum, the substantive domain and projected performance levels for in-program objectives and, where feasible, the magnitude of post-program impacts upon trainee performance and organizational functioning.	To design and manage a program that promotes reaching objectives in an accountable manner, the achievement of objectives should be formulated in demonstrable terms.
Step 5	Design a program serving program objectives.	The design details how and with what resources the performance of tasks identified in program objectives is to be systematically brought about.
Step 6	Pretest trainees and/or their departments to determine pre-training performance levels.	Pretesting provides a comparative basis for potential evaluations of in-program trainee achievement, trainee on-the-job performance, and organizational functioning, and gives program staff advance information about trainees or their departments so the program can be tailored to address current trainee needs.
Step 7	Conduct the program in a manner that corresponds with program design and addresses actual trainee needs.	For program activities to address specified objectives and for intended performance changes to be demonstrated, sufficient controls must be placed on the program design that it is implemented as planned, yet flexibly enough to accomodate the immediate needs of the current trainee audience.
Step 8	Evaluate the in-program effects of training on participants.	Even where it is difficult to define how the program is linked to post-program outcomes, the measurement of those in-program outcomes under the program's control is often feasible, can provide useful information about program effectiveness, and suggests modifications needed in objective setting, design, or implementation.
Step 9	Evaluate the effects of training on the work setting in terms of trainee performance and/or organizational functioning.	Where the links between in-program outcomes and later trainee performance and/or organizational functioning are sufficiently well-defined, or where trainees have at least clearly articulated their own perceived performance deficits, the measurement of training's impact on the work setting offers the strongest evidence of a program's goal-attainment.
Step 10	Use evaluation results in subsequent program development and revision.	Evaluation results can not only document the extent of program effectiveness, but also suggest ways in which the program or its environment need to be changed.

In the discussion that follows, we take each of these 10 core steps, break it down into substeps, articulate a rationale for each substep, and describe current practice in relation to the substep. The evidence clearly fails to justify the comment by one POST administrator that "training is pulled off the wall everywhere." Still, it shows that current realities in police management training often stand in sharp contrast to the structured, highly deliberate, phased process derived from industry. These findings can be used to assess how closely a particular program or set of programs follows the industrial model and to identify new approaches to program development.

A. STEP ONE: SETTING TRAINING GOALS BASED ON BALANCED INPUTS

The first developmental step is to "set broad training goals based on balanced inputs from those to be affected by a proposed program." This step may be further broken down into the four substeps in Exhibit 4.

Exhibit 4
Setting Training Goals Based on Balanced Inputs

Substep	Rationale
Obtain balanced inputs from those with a potential interest in the program's operations and outcomes toward the establishment of program goals.	Individuals and groups affected by program development, delivery, and impact can provide relevant information about current performance expectations, perceptions of training needs, and means for addressing needs. Interested parties can include both the direct consumers of training in law enforcement and the indirect consumers of training in other criminal justice agencies, community organizations, and the general public.
Synthesize goal inputs, determining their priorities and potentially contradictory emphases.	Synthesis of goal inputs can be necessary to capture divergence in emphasis, priority, and perspective among the contributing groups.
Set training goals, specifying the conceptual thrust and scale of the proposed program.	Goals provide general direction for program design and a specific framework for assessing needs.
Circulate proposed goals among parties that provided input to obtain feedback and secure commitment.	Circulation of proposed goals for feedback and commitment ensures that goals adequately reflect the range of interests that contributed to goal setting and helps maintain continued support.

1. Obtaining Balanced Inputs From Affected Groups. Program developers have made few documented efforts to draw on diverse criminal justice and community agencies in establishing initial program goals. Our national survey showed that only four sets of individuals exert strong influence on program goals. In declining order of strength, they are: program staff, program instructors, POSTS, and police chiefs and executives from user agencies. Four other groups have only moderate influence on goals. Again, in descending order, they are: training directors from user agencies, program advisory board members, graduates of related programs, and local academic institutions. Everyone else--including state legislators and local government officials, national and state police associations, police unions, SPAs, citizens' groups, and public interest groups--seems to contribute little or nothing.

Based on the mail survey and on-site observations, it seems that program developers often either presume program goals, based on their own intimate experience with training needs, or carry goals over directly from existing or prior programs. If they solicit input from law enforcement agencies, they generally do so informally, through casual conversation and observation of operations. Informal input is sought mostly from senior departmental officials and training officers but rarely from all ranks. If sought at all, formal input is obtained initially and periodically thereafter from the executives of primary user audiences in group settings; for example, through annual professional meetings, executive training sessions, and meetings of the sponsoring organization's advisory board or of a professional group's education and training committee. Formal input is also sometimes sought from trainees taking part in related programs through questions in a reaction survey about "need for significantly different types of programs" or from instructors in the same programs. In most cases, however, even where input is systematically sought from directly affected groups in law enforcement, the program's goals have already been largely determined, and those contributing apparent goal inputs are generally asked to focus on specific topical and content needs rather than on the program's overall scope and scale. Their periodic inputs serve more to reaffirm goals than to create them. In some exceptional cases, program developers have obtained goal inputs from other criminal justice agencies, state legislatures, local government agencies, and the business and academic communities from which instructional assistance will, in all likelihood, later be sought.

2. Synthesizing Goal Inputs. To the extent that the program developer obtains real goal inputs, the later determination of their priority and potentially contradictory emphases is generally accomplished informally and is not documented. Because program developers typically obtain goal inputs from relatively restricted user groups, prioritization and synthesis are not usually called for this early in program development. It becomes more important later, in the needs assessment process. There are clear exceptions to this rule, but the synthesis of divergent goal inputs leads more likely to a "decision not to decide," i.e., to couch divergent notions of the desirable direction for the program in neutral language that masks differences among program models. On a system rather than a program level, the FBI has long lived with shifting goal priorities in the programs it has conducted for the field. From time to time, the Bureau has placed primary emphasis on three areas: the broad development of managers in response to immediate demonstrated trainee need, response to all

agency requests for training, and compliance with agency requests where capacity for and interest in training self-sufficiency have been first demonstrated. Only recently has the Bureau stated an interest in integrating the three.

3. Setting Training Goals. Whether based on systematic inputs from affected agencies or on administrative fiat, each program ends up with certain goals, some explicit and others implicit. These goals generally specify the scale of a program--week-long traveling management schools, an eight-hour department-based program, a three-week residential program, a career development sequence involving each officer in 40 hours of in-service training per annum, and so on--more clearly than a program's conceptual thrust, which is more subject to debate. The conceptual thrust of a program is generally articulated in terms that all of the diverse projected (but probably not consulted) user groups will find palatable. Most programs express their scope largely in terms of "exposure to" a set of concepts, divorced from explicit notions of changes in police agencies. In contrast, the LEAA-funded Police Executive Institute, the originally LEAA-funded Police Executive Development Institute (POLEX) at Penn State, and some elements in the FBI all explicitly define their purpose as moving police agencies away from an authoritarian and toward a more participative model of organizational functioning. Even in these examples, it seems that a clear definition of program goals evolved from experience with the program rather than guided program development.

4. Circulating Goals Among Those Providing Goal Inputs. When determination of program goals is largely a function of departmental command orientation and executive decision, the circulation of goals through the chain of command is intended to ensure commitment, not to obtain feedback. In programs serving multiple agencies, the circulation of program goals, however established, is typically used to probe affected groups, especially chief executives, about the precise dimensions of training need. Except in isolated cases where goals are formally presented to an advisory board or similar body for review, the circulation of program goals presumes general commitment and is designed to trigger or continue the task of needs assessment.

B. STEP TWO: CONDUCTING AN ASSESSMENT OF TRAINING NEEDS

The second developmental step is to "conduct a needs assessment to refine the dimensions where discrepancies exist between current and desired levels of individual performance and/or organizational functioning." This step may be further broken down into the two substeps in Exhibit 5.

Exhibit 5 Conducting an Assessment of Training Needs	
Substep	Rationale
Specify the scope of needs to be assessed in terms of: characteristics of the target population, organizational vs. individual emphases, exact functional and other areas of interest, and standards against which individual and/or organizational performance might be assessed.	Before formulating a needs assessment plan, broad program goals must first be focused upon specific functional and other areas of individual or organizational practice for which standards can be specified and current performance levels ascertained for a given target population or organization.
Develop and implement a needs assessment plan employing techniques that are both appropriate and feasible within resource constraints.	To isolate actual performance deficits and provide relevant information for setting program objectives, needs assessment techniques must reliably measure salient aspects of current and desired performance.

1. Specifying The Scope Of Needs To Be Assessed. In preparing to collect information that will elucidate those areas where current individual performance or organizational functioning is inadequate, most program developers begin with a clear notion of the level and general responsibilities of the intended target population. Beyond this, it is often unclear whether the needs assessment is to focus on individual development or organizational change. Where individual vs. organizational emphasis has been determined, little further indication is given regarding the exact functional or other areas of interest. Except in a limited number of compliance-oriented programs, program developers do not identify standards against which individual and/or organizational performance might be assessed. This stems largely from program focus on broad individual development, as will be reaffirmed throughout this chapter.

2. Developing And Implementing The Needs Assessment Plan. The term "needs assessment" refers not to one but to myriad approaches for clarifying the dimensions of training need for the target population from which trainees will be drawn.[4] Because many have come to regard this catch-all term as a meaningless

[4] Needs assessment is the most hotly debated issue in program development. Analysis of needs assessment models fell outside the scope of this study. Two ongoing LEAA-funded studies deal directly with needs assessment. Gil Skinner at Michigan State University is developing a series of manuals on program development and needs assessment, and Travis Northcutt at the University of South Florida is analyzing alternate approaches to training needs assessment. In addition, in Chapter Six, we explain how to use a standardized instrument, called the Managerial Training Needs Profile, to design a course, select participants, and/or evaluate results.

"buzz word," one must address with caution such a question as "Was there a needs assessment?"

A wide range of activities pass as "needs assessments." Some are formal, many are informal. Some are directed to the target audience, many are more diffuse. Our national survey showed that program developers rely strongly on only two techniques "to clarify the nature of the performance deficiencies giving rise to the need for training." These are: informal interviews with chiefs and other user group executives and informal interviews with program graduates. They rely with moderate frequency on four other techniques. In declining order of use, they are: posttraining reaction surveys of trainees from prior sessions, formal surveys of target audience incumbents, job-task analysis, and formal surveys of chiefs and other user group executives. Relatively little use is made of individual testing of the target audience, formal surveys of entire agencies, management audits of participating agencies, and delphic surveys of experts.

Observations of operating programs generally support our national survey. Prior to program design, training needs are assessed largely through informal conversation with chief executives about departmental needs and with former trainees about the usefulness of prior training. Where this informal process takes place within a single police agency, the heads of operational divisions are also likely to be consulted. If formal surveys of incumbents or user group executives are conducted, they typically solicit topical recommendations for the consideration of programs serving multiple agencies. The more complex topical surveys identify not only topics of interest but also the prioritization of topics and desired subject matter content. The periodic topical survey, however, does not attempt to pinpoint performance deficiencies; instead, it moves right to the task of curriculum construction.

The only apparent discrepancy between our national survey and site visit data pertains to job-task analysis, and this difference is easily explained. The national survey seemed to indicate that program developers use job-task analyses with moderate frequency, as often as they formally survey target audience incumbents and more often than they formally survey chiefs and user group executives. By contrast, other sources suggest that, even where job-task analyses have been employed in the development of a jurisdiction's basic recruit program, this technique has rarely been extended to management training. How can we explain this discrepancy? First, job-task analysis is fast coming to be seen as the "right" way to assess training needs. Some survey respondents may have overreported their use of job-task analysis to "look good." Second, many program developers may have borrowed an already existing job-task analysis from another program, perhaps one in another jurisdiction.

Instead of examining the performance deficiencies or topical interests of the target population, the program developer often borrows related information from other jurisdictions in lieu of doing a formal needs assessment. Our national survey showed that program developers rely heavily on two existing information sources to determine training needs. These are: needs assessments developed for prior programs in the same jurisdiction and course specifications and curricula outlines from recognized programs. We might question what they mean by

"needs assessments developed by prior programs," given what often passes for needs assessment. With moderate frequency, they also use the following sources (listed in declining order of importance): centralized data on the training histories of the target population, studies of police manager effectiveness (other than the National Manpower Survey), and national standards on police manager needs (including the Police Executive Report). With low frequency, they use job-task analyses borrowed from other jurisdictions and other forms of borrowed needs assessments. They make almost no use of and seem to attribute little value to the National Manpower Survey. Typically, the program developer weighs any or all of this borrowed information along with informal discussions with chiefs and program graduates and moves directly to curriculum development without clearly stating performance deficiencies and articulating objectives.

In conclusion, little empirical needs assessment takes place before a program is designed. The type of needs assessment typically used hardly begins to tap the perceptions of target population incumbents about deficiencies in performance and organizational functioning, tends to overlook changing concepts of the police manager's role and of police organization, and is too sensitive to what police managers want rather than what they need. It moves too rapidly from identification of topical interests to curriculum development without first locating gaps between current and desired performance, determining whether these deficiencies are really correctable through training, setting objectives to help fill these gaps, and then establishing a curriculum to meet objectives.

C. STEP THREE: SETTING TRAINING OBJECTIVES

The third developmental step is to "develop training objectives to meet identified needs in the context of overall program goals." This step may be further broken down into the four substeps in Exhibit 6.

Exhibit 6
Setting Training Objectives

Substep	Rationale
Synthesize the results of needs assessment with the original inputs to goal setting to provide the basis for the establishment of training objectives.	Although needs assessment results offer an empirical basis for setting objectives, they should not be allowed to dominate objective setting because their emphases are largely a by-product of the techniques used. The results of one or more needs assessments should be synthesized, therefore, with the original inputs to goal setting.
From this synthesis, identify deficiencies in individual performance or organizational functioning that can be addressed effectively through training, rather than some other means.	Some performance deficits may not be amenable to remediation through training but may be addressed through other means, such as job enrichment or improved procedures for assessing promotion potential. Additional experience may be the main remedy for performance deficits of some types.
Formulate in-program objectives in terms of outcomes that are under the program staff's control and are, hence, plausible.	The identification of plausible objectives to be attained before the completion of a course allows efficient mobilization of resources and energies toward outcomes minimally affected by external factors.
Indicate how these in-program objectives are presumed to be linked to longer-term objectives involving on-the-job performance and larger organizational impacts.	Indicating how attainment of in-program objectives relates to longer-term objectives ties the program's activities and objectives back into the world of work, showing the extent to which changes are seen as desirable and the ways that in-program outcomes link with performance changes.

1. Synthesizing Needs Assessment Results With Goal Inputs. Because program developers often obtain goal inputs and needs assessment data at the same time, both through a prematurely administered topical interests survey, this information need not be synthesized. There are cases where this synthesis is warranted and can occur, however. The initial executive-sanctioned goals for one Houston PD program called for a day of activities to "correct bad habits." Before he designed the program, the training director conducted a survey of all ranks to determine their common perceptions of a first-line supervisor's training needs. He then synthesized these with an existing departmental task analysis. When the results of these synthesized needs assessments showed the need for a refresher/update program, the original scale of the program was expanded and the corrective and update purposes for training were merged.

Programs operating as satellites of a statewide POST-certified program represent a special case in goals/needs synthesis. POST-dictated curriculum requirements, derived from needs assessments of varying soundness, impose certain obligations on offering agencies. These vary from general compliance with descriptive titles and broad curriculum outlines to adherence to particular terminal performance objectives and instructional techniques. Within the framework of these requirements and certain given goals, programs usually retain some flexibility to assess needs locally and follow through on the local needs assessment with construction of additional program elements.

It is more often necessary to synthesize several types of information about training needs than to synthesize goals and needs. Typically, the program developer is in the position of reconciling informal conversations, topical surveys directed to several affected groups, and other information borrowed from outside the jurisdiction. As the Houston example of goals/needs synthesis shows, the results of formal needs assessments can also be integrated at this point. In practice, in integrating either diverse needs assessments or needs assessments with goal inputs, a holistic approach is rarely taken. Instead, the results and implications of one needs assessment are either adopted in entirety or are altered radically by the inputs of one affected group.

2. Identifying Deficiencies Remediable Through Training. We can independently treat two questions here: Are specific deficiencies identified? Are identified deficiencies that are remediable through training separated from those that are not? The first question has already been answered. Much more often than not, programs specify broad topical interests rather than particular deficiencies in individual performance or organizational functioning.

The second question, whether identified deficiencies that are remediable through training are separated from those that are not, is more complex. If a program's primary stated purpose is "exposure to and transmittal of a body of knowledge," as is often the case, then remediation of performance deficits has not arisen as an issue. Any presumed deficits have not been spelled out and remain implicit. There may be variation among program operators in what they see as "the problem." Most programs build on diffuse notions of which individual and organizational behaviors even require change, so program operators rightly presume the need for a lengthy postprogram experience in testing and determining the validity of knowledge obtained in training. This postprogram experience they intend to be "qualitatively different" from the traditional experience of the police manager who fishladdered through the ranks without benefit of exposure to a specific body of managerial knowledge.

If we can tease out how program operators intend this body of knowledge to impact on job performance and organizational functioning, it seems that they view training as one among many means for upgrading police managers and the management of police agencies. During site visits, they also took the position, implicitly or explicitly, that most police agencies, even while making training a low priority, have traditionally relied on it too heavily. Most agencies have expected from training what might be more effectively accomplished through other means. These include better procedures for assessment and promotion of managers, job exchanges, job enrichment, and several forms of internships.

3. Formulating Plausible In-Program Objectives. It is a judgment call whether the objectives that are set for accomplishment within a course are under staff control and are, hence, really plausible. To assess the plausibility of a program's in-program objectives, one has to consider both explicit and implicit objectives, independently and in interaction with each other; the amount of resources brought to the program; the effectiveness with which resources are mobilized; and the relationship between expectations and confirmed general theory.

Most program managers and operators do not make explicit all the objectives they intend to accomplish, and many of those explicitly stated are still unclear. The most explicit and clear objectives generally refer to transmittal of a body of knowledge. When program objectives refer to transmittal of a body of knowledge and development of skills for using it, these objectives tend to be under the program's control. In part, this depends on the instructional staff's ability to accommodate the immediate trainee audience's needs. Objectives become progressively less plausible when the program operator has not clearly delineated the body of knowledge to be transmitted, when instructors fail to cover material or neglect to reinforce each other's presentations, or, in the extreme case, when instructors contradict other presentations without attempting to acknowledge differences. Efforts to transmit a body of knowledge can, thus, be implausible because its content is ill-defined, because resources are inadequate, or because resources are ineffectively brought to the task.

Objectives dealing with attitude change, which tend to be left implicit, are generally less under the control of program staff and are, hence, less plausible than objectives related to knowledge transmittal. Attitudinal objectives point in many directions. They might involve changes in attitude toward:

- The manager's role in motivating subordinates
- The value of systematizing behavior according to business principles
- The usefulness of research data in upgrading practice
- The effectiveness of fitting one's management style to the situation at hand
- The need for making management systems in police agencies more participative
- The value of providing greater opportunties for self-actualization on the job
- The importance of reducing the insularity of police departments from other community agencies and of police executives from each other

The intensity and duration of most programs, however, are insufficient to "shake" or "crack" deeply entrenched attitudes, much less to change them radically. The typical program can control transmittal of knowledge and acquisition of skills (or tools) for using this knowledge. It can also control whether certain structured experiences, geared to reinforce knowledge and trigger attitude change, take place and whether trainees can attach appropriate explanatory concepts to these experiences. But, except where a program contains a mechanism for prolonging the training experience beyond the single course--either through a network of program graduates or a structured career sequence--significant in-program attitude change is rarely a reasonable expectation.

Some programs, however, view attitude change as indispensable to maintain the process of molding a police manager. Some program operators recommend that such programs aim at generating in trainees a state of cognitive dissonance, i.e.,

a state of discomfort caused by the incompatibility of current beliefs and new knowledge, the validity of which has been partly demonstrated in training. Within all but the longest and best-integrated programs, it is not reasonable to expect more attitude change than found in this initial state of cognitive dissonance. Several program instructors said, "We want them to leave here confused, uncomfortable with their old beliefs and with what they have been doing, but not yet comfortable with what we are saying either." The literature in social and experimental psychology shows that this state of discomfort naturally leads to efforts to decipher the sources of dissonance and to test the validity of new knowledge in the work place. This leads to the dissipation of dissonance and to gradual attitude change if the work experiences validate the new knowledge.

4. Linking In-Program Objectives With Longer-Term Objectives. Most program operators say they cannot pinpoint how they want in-program outcomes to affect job behavior and the work setting. Those focused on compliance with departmental policy are the only possible exception. Programs vary in how precisely they try to relate in-program outcomes to later performance and organizational functioning. Those with a clear and plausible conception of the results they want to see within the program are usually better able to define what can be expected to follow once the trainee returns to the job. In coming to this conclusion, we do not equate "clear in-program objectives" with "terminal performance objectives" because one does not imply the other. Objectives can be clear, even if empirical success criteria have not been identified. In contrast, quantitative criteria can be used to disguise a program's lack of a unifying theme under the mask of empiricism.

In most programs, instructors differed not only in their opinions of what trainees should know but also in the range of performance changes and impacts on the organization that they found acceptable. In many programs, more than one logically consistent set of expectations seemed to be in play simultaneously, as though two partially articulated programs existed side by side. When several models were partially integrated in this fashion, in-program objectives were not clearly tied to longer-term objectives. Rather, the expectations of one seemed to neutralize or overshadow those of the others. This may be simply another way of saying that clear and plausible in-program objectives, united by a central theme, may be the precondition to adequate definition of how training's outcomes will affect the world of work.

D. STEP FOUR: TRANSLATING OBJECTIVES INTO SUCCESS CRITERIA

The fourth developmental step is to "translate objectives into measurable success criteria reflecting, at a minimum, substantive domain and projected performance levels of in-program objectives and, where feasible, the magnitude of post-program impacts upon trainee performance and organizational functioning." This step may be further broken down into the three substeps in Exhibit 7.

Exhibit 7
Translating Objectives into Success Criteria

Substep	Rationale
Formulate program objectives in terms of trainee learning and behaviors that are amenable to measurement.	For the accomplishment of program objectives to be demonstrable, program objectives should first be broadly amenable to measurement, even before attempts are made to view them in terms of discrete sets of criteria.
Identify acceptable empirical criteria to indicate the attainment of in-program objectives and, where feasible, related post-program objectives.	To demonstrate performance changes, measures must be identified that can reliably indicate intended changes. Identification of empirical criteria defines the conditions under which program effects can be observed, documented, and verified.
Identify multiple complementary criteria where appropriate and feasible, especially for "softer" objectives not readily amenable to measurement.	A single type of criterion measure can only partially indicate the attainment of an objective. Multiple complementary criteria provide convergent validity and increase the extent to which program results can be demonstrated.

1. Formulating Objectives Amenable To Measurement. Objectives that involve transmittal of knowledge and development of skills for using it tend to be amenable to measurement when their content and limits are clearly detailed. In contrast, objectives calling for attitude change do not readily lend themselves to measurement. This is so even when their accomplishment is more plausible, as in long-term programs and those with other means for prolonging the training experience. Efforts to document attitude change have typically employed attitudinal measures that related poorly to program objectives. In part, this is because appropriate instrumentation is difficult to locate or to construct.

2. Identifying Empirical Success Criteria. Most programs do not identify empirical criteria for measuring attainment of objectives. Based on our national survey and site visits, about two programs out of five have at least tried to identify empirical success criteria. Among the agencies that we observed, the Kentucky and Oregon POST Councils, the Northwestern Traffic Institute, the Oakland PD and the District of Columbia PD (with FBI assistance) all had identified "terminal performance objectives" for some of their courses. These criteria set forth knowledge or behaviors to be demonstrated during or at the completion (or terminus) of training as evidence that particular objectives were attained. In some cases, however, the instructors could identify the criteria but did not

actually use them to guide instruction or to measure outcomes. Sometimes curricula had significantly changed but criteria remained the same, suggesting that they served as part of a contractual agreement or fulfilled other requirements. Even when success criteria were developed and then not systematically used, however, their dissemination among program staff seemed to increase consistency in program implementation.

3. Identifying Complementary Criteria For "Softer" Objectives. Of those programs that employed terminal performance objectives, only experiential programs measured the more abstract objectives in more than one way. Structured experiential exercises used in the assimilation of conceptual material were measured two ways: paper-and-pencil testing of knowledge and successful completion of a structured experience or other simulation exercise demonstrating that this knowledge was assimilated. Some nonexperiential programs also used behavioral demonstration of newly acquired skill as a complementary measure of knowledge transfer. The conceptual nature of this material, however, generally impedes use of skill demonstration as extensively as in operations-oriented programs.

E. STEP FIVE: DESIGNING A PROGRAM TO SERVE OBJECTIVES

The fifth developmental step is to "design a program serving program objectives." This step may be further broken down into the seven substeps in Exhibit 8.

Exhibit 8
Designing a Program to Serve Objectives

Substep	Rationale
Specify the learning principles upon which the program will be based and the corresponding instructional approaches for satisfying program objectives.	Programs vary widely in substantive emphases and general orientation. Learning principles appropriate to one program may fit poorly in another. Specification of learning principles and instructional approaches focuses program design on the chosen objectives.
Develop a curriculum that corresponds in content and instructional technique with program objectives and that specifies the amounts and types of information to be presented.	By detailing substantive content and instructional techniques, the curriculum provides a framework that guides the instructional process toward program objectives.
Establish trainee selection standards in line with the training needs to be addressed by the program.	Training resources can be more effectively directed toward a well-defined trainee audience demonstrably in need of training and capable of attaining the performance levels specified in program objectives.

Substep	Rationale
Establish trainer employment standards that reflect subject matter familiarity, training capability, compatibility with program philosophy, and trainer-trainee rapport.	Trainers constitute the main vehicle for communicating program contents and directing program activities. Hence, it is important that trainer employment standards require both individual capabilities and readiness to mesh with the program.
Develop evaluation instruments that measure plausible outcomes as part of program design.	Although evaluation instruments may have to be modified to reflect differences between the planned and actual programs, the overall evaluation strategy has implications for program operation. Thus, it should be an integral part of design and not an afterthought or the artifact of funding requirements.
Coordinate program design with any similar programs offered by the same organization in related topic areas.	Coordination with similar programs under the same auspices helps to allocate training resources effectively. It reduces duplication and builds bridges among programs.
When contracting to an outside organization to conduct training, ensure that their design meets program goals and objectives.	Regardless of who designs a program, if it does not build on soundly rooted program objectives, the effort is unlikely to produce systematic effects.

1. Specifying Learning Principles. Most programs rely mainly on traditional lecture and discussion methods without clearly stating the learning principles, or conceptual rationales, for the instructional techniques that they use. The visited programs that specified learning principles usually focused on the importance of involving the adult learner actively in the learning process. Two such programs used the term "androgogy," coined by Knowles, in reference to the science of adult education. Programs that follow this active learning approach operate on the principle that involvement in learning leads more rapidly to internalization of knowledge and to the assimilation of behaviors into one's repertoire. Among these programs were the Oregon POST Council, the Police Executive Institute, the Oakland PD, and POLEX. To involve the adult learner, these programs specified the importance of programmed materials that sensitize the trainee to concepts before they are formally introduced, written active exercises, structured group experiences, and other manipulated trainee interactions. The contractor serving the Oregon POST Council expressed the principle on which most of these programs build, saying, "If a picture is worth a thousand words, an experience is worth a thousand pictures." A fifth program specified an organizational development model. This model called for confronting trainees with the

ideal system, then gradually moving them toward this ideal and eroding the status quo's validity by analyzing differences between current practice and the ideal and the advantages/disadvantages of each. This was capped by development of an implementation plan, which helped commit the trainees to trying new ideas. For the most part, however, programs do not specify learning principles behind their instructional approaches.

2. Developing A Curriculum That Reflects Objectives. Do the curricula that programs adopt really correspond with their objectives? This was among our major interests on site visits. Initially, we started to examine three related questions: Does the substantive content of curricula correspond with objectives? Do curricula specify instructional techniques appropriate to these objectives? Do curricula describe the amount and type of information to be presented in enough detail to give meaningful direction to instructors? When we ran into difficulty with our original questions--most programs do not clearly state objectives--we had to shift our focus slightly. Ultimately, we asked an altogether different question: not "Do curricula correspond with objectives?" but "How much fluidity in curricula and objectives is desirable?"

It quickly became futile to ask if curricula reflect objectives because most programs do not consistently state objectives clearly. With the general exception of programs that employ terminal performance objectives--showing, in measurable terms, the areas in which change is desired and the intended extent of change--most programs do not specify in-program objectives. To guide the instructional process, nearly all programs provide a general outline of curriculum topics. Two out of three, based on the national survey, have broadly articulated course objectives. About two in five provide instructors with module-by-module course content summaries and assume that objectives will be self-evident. A like number, two in five, use terminal performance objectives. When course outlines and objectives do not detail intended instructional methods and course content, however, course objectives tend to become whatever the available instructors choose to make them. This frequent obscurity in objectives makes the basic issue--the correspondence of content and method with objectives--almost moot.

By slightly shifting the focus of our initial questions, we can still deal with the issue, although in a roundabout way. Why could we not infer objectives from the content and methods in the curriculum outline and then use these inferred objectives to answer our new questions? Admittedly, this sounds like circular reasoning, but it is not if we phrase our new questions carefully. Our new questions might be: Are course outlines sufficiently detailed to allow the inference of course objectives? Is there internal consistency in content? Are instructional methods appropriate to content? Because programs lack consistently clear objectives, this new set of questions is more appropriate.

Our visits to programs strongly suggested that, if a program sets clear objectives that are held together by a unifying/central theme or themes, then its curricula usually specify the content and method that are appropriate to the objectives. It enumerates the amount and types of information to be presented in enough detail to guide instruction. By contrast, where a program's objectives are unclear and do not center around a unifying theme, it is difficult or impossible to infer objectives from course outlines. Content appears internally inconsistent, and almost no mention is given of instructional methods. Trainees and instructors alike often referred to this type of curriculum as a "vegetable

garden" or "smorgasbord" approach to training because there were no clear objectives or unifying themes. There are exceptions to these generalizations, however. Even programs with clear objectives often state no instructional method or allow for a method that may be inconsistent with objectives. This discrepancy is especially obvious in programs that explicitly aim to familiarize trainees with participative management theory and principles and implicitly attempt to shift values toward more participative management philosophy yet, in their curricula, do not call for methods that encourage trainee participation.

In practice, both objectives and curricula have some fluidity. How much fluidity is desirable? The foregoing discussion might convey the impression that objectives should be etched in stone. Experience with program operation, however, shows that programs evolve, largely on the basis of trial and error, in response to changing conceptions of need. A later core step will address the importance of making changes in curricula, trainee selection requirements, trainer hiring standards, objectives, and so on, based on program evaluations. Though it may appear premature at this point to stress the value of sharpening objectives and curricula through experience with program delivery and sensitivity to changing needs, it is essential for understanding issues of design and implementation. As cases in point, we cite two observed programs.

The POLEX program at Penn State began in 1970 with a broad concept of the type of program desired, but the initial curriculum, according to current staff, "took the shape of the faculty available from session to session." Faculty were initially chosen for one of two reasons: They were recommended by a municipal PD training director or they had high immediate renown or visibility. With experience and with help from Penn State's own management development faculty, POLEX developed a central/unifying theme and wove its threads through the curricula and objectives, dropping elements that did not fit or support it. The resultant clarity in its objectives allows POLEX the flexibility to modify program components to serve trainee needs in the context of overall program goals.

A second program, in the throes of reorganization when visited, illustrates the tribulations a program can face in the attempt to develop a unifying theme, state objectives clearly, and formulate a curriculum corresponding to these objectives. In this program, objectives remained largely the province of each individual instructor, except that an informal counterprogram seemed to be emerging through the unofficial cooperation among a few of them. This counterprogram aimed at clarifying objectives and clearly demarcating the areas of concentration to be covered by each instructor in a manner that would allow each to build on the other's presentations and, thus, generate a real course progression. To change the program, these instructors helped "set up" other instructors to disqualify them and show the inadequacy of their presentations. Sensing this attack, these other instructors used defensive diversions to prove they were the "policeman's friend." Many trainees sensed this and helped subvert some presentations in sympathy for the counterprogram. Meanwhile, the program coordinator was independently trying to develop a system of terminal performance objectives. In this second program, instructors seemed to have more freedom than was desirable. There was too much fluidity in curricula and objectives because neither was clearly stated around a unifying theme.

Most program developers agree that curricula should be designed to serve objectives. They also agree that curricula should be modified to reflect changes in objectives brought on by sensitivity to evolving needs. Yet curricula have typically been generated through consensual development. As one POST administrator noted, "People sit around a big table and play a numbers and titles game. They never get to focusing on job content. The system is crazy. Curricula are developed by the seat of the pants." And often these curricula, on or off the mark when formulated, become "etched in stone" and continue to provide the framework for instruction long after their use has become counterproductive.

3. Establishing Relevant Trainee Selection Standards. Programs use a wide range of criteria in trainee selection decisions. These criteria tend to cluster into several distinct selection models. It is also clear, however, that there are certain criteria that have relatively stronger or weaker influence on trainee selection in all programs.

When asked on the national survey about the extent to which certain criteria are important in trainee selection, program managers and operators rated only two criteria as highly important: current responsibilities and rank. Two criteria are of moderate importance: recommendations by supervisors and individual's demonstrated training needs. Six other criteria are of low importance. In declining order of importance, these are: other prior training, "promise" for higher levels of responsibility, time elapsed since last participation in a career sequence, time in grade, scores on promotional exams, and prior education. Many criteria that one might expect to be relevant to selection decisions are not considered important.

Our on-site observations of training help explain why certain selection criteria should not be expected to have uniform relevance to all programs. There seem to be four distinct models of trainee selection. They reflect the varying auspices and financial arrangements under which programs are conducted and the divergent purposes behind training. The four models are: pre-service and other rank-related selection, career progression selection, departmental discretion selection, and status selection.

- Preservice and other rank-related selection. Found in programs offered under the auspices of a single department, criteria primarily reflect scores on promotional exams and time in grade. Criteria tend to be strictly enforced by a department's training division.

- Career progression selection. Ordinarily found in programs certified by a POST Council and offered under varying auspices, main criteria are rank, responsibility, and completion of prior installments of a defined career sequence. Other criteria might include: prior education, other prior training, time in grade, and time elapsed since last participation in the sequence. Criteria are generally prioritized, and the offering agency has flexibility in obtaining "waivers" of criteria to fill a class.

- Departmental discretion selection. Ordinarily found in an independent program based at an academic institution, criteria state the rank and other characteristics of the person most appropriate for the program. It is up to the department to use its discretion in selecting those who warrant training. Designated trainees are generally accepted by the program on a first-come, first-served basis. Field programs conducted by the FBI National Academy's Management Unit, including the Management Schools, also tend to fit this selection model best. This is because the Bureau recommends to the host agency the appropriate rank structure for the class but does not define the term "police" and leaves responsibility for ultimate selection up to the host, which is usually a regional academy.

- Status selection. Found in only a few select organizations, criteria reflect the immediate status or projected promise of the participants, who are to be "creamed" from a pool of applicants. Criteria include holding the chief or other senior executive position in a major law enforcement agency, recommendation by prior participants and immediate superiors, an untainted record as shown by background investigation, and correspondence with the desired geographical mix for a given class. Programs using a combination of these and other criteria include: the FBI National Academy, the FBI National Executive Institute, the Northwestern Traffic Institute's "long course," the Police Foundation's Police Executive Institute, and the National Sheriff's Association's National Sheriff's Institute.

Nearly all programs use one of the four selection models above, with some variation to reflect the particular program's purposes. Seen from the perspective of the trainee whose department has to mesh its own needs with the selection standards of available programs, selection criteria are rarely made clear. A trainee ordinarily has no way of determining whether his selection for training came about as a reward or sanction, in recognition of his promise, or because he is in need of correction and a "kick in the pants." The criteria for selection decisions remain implicit at best. Many selection decisions seem to be made oblivious to relevant facts. Thus, the trainee often leaves asking, "Why me?" On two different levels, those of the training program and the sending agency, trainee selection standards rarely focus on the documented needs of the individual trainee or on the trainee's later prospects for implementing training. To meet training needs better by effectively matching trainees with available programs and clarifying for trainees what is expected of them, the POST Councils in California and New Jersey developed or are developing a departmental training assessment manual. In most states, however, there is no mechanism for helping all interested parties to concur on why a given trainee has been selected for training.

4. Establishing Relevant Trainer Employment Standards. Trainer employment criteria are as diverse as training's auspices and purposes. As in trainee selection, certain criteria have relatively stronger or weaker influence in trainer hiring decisions in all programs. Based on our national survey, the consideration of paramount importance is experience directly related to the subject area for x years. Six other criteria are also rated highly important in trainee selection. In declining order of importance, these are: law enforcement experience, sincere interest in teaching police officers, experience as a teacher in the subject area, advanced academic achievement, submission of a sample lesson plan, and completion of an instructor's course. Six other criteria are rated of

low importance. In declining order of importance, these are: congruence with the program's philosophy, as shown in oral interview; immediate rank in department; personal recommendation by another instructor; university teaching experience; national recognition on the subject to be presented; and graduation from the program. From this evidence alone, it appears that trainer hiring criteria fall evenly into two categories: high importance and low importance. Some of those rated highly important tend to be procedural (e.g., submission of a sample lesson plan), wherease some rated of low importance appear more relevant to hiring decisions (e.g., congruence with the program's philosophy as shown in oral interview).

Our on-site observations, again, help amplify the survey data. The criteria that are rated highly important tend to be supported by official policy. The criteria rated relatively unimportant are rarely backed by policy but, in many cases, are more important in making the actual hiring decision than the highly important ones. Personal recommendation by another instructor or graduation from the program are often subtle, nonpublic hiring criteria.

The employment criteria for most visited programs recognize that the appropriate credentials for instructing police managers in management cannot be stated unequivocally and depend on the program's exact purposes. Most programs allow for a mix of backgrounds. Relatively few absolutely demand law enforcement experience. Because of either uncertainty about appropriate credentials or recognition of the diverse combinations of experience that can make an effective instructor, most programs permit compensating criteria, such as:

- Advanced academic achievement or completion of high school and an instructor's course
- Specified years of experience in law enforcement or an equal period of experience in a field directly related to the topic
- Advanced academic achievement and specified years in law enforcement or graduation from the program
- Current departmental rank or advanced academic achievement plus specified years of experience in a directly related field

Trainer hiring standards partially mirror the trainee selection process and program auspices. For example, departments generally require current departmental rank. Academic institutions tend to demand academic credentials. Status programs tend to require their instructors to have "timely notoriety" or national renown. POST-operated programs seem to allow the greatest flexibility in compensating one experience for another.

So, do programs use relevant standards in trainer hiring? Most programs do not explicitly call for compatibility with program philosophy. This may be because many program operators do not recognize that there is a multitude of management training philosophies. Most programs do take into account subject matter familiarity, training capability, and trainer-trainee rapport. Programs that require law enforcement backgrounds do so on the assumption that only those who have "been there" (i.e., served as a police officer) can relate to and be accepted by police trainees. This assumption is congruent with certain training

models but has been sharply contested by some programs. Before hiring a trainer, however, few programs try to assess subject matter familiarity, training capability, or trainer-trainee rapport directly.

5. Developing Evaluation Strategies As Part Of Program Design. Program developers seem to give little thought at the time of program design to construction of evaluation strategies. The general exception is programs using terminal performance objectives, which often develop program examinations in accord with success criteria during program design. Programs that have tried to measure long-term impact on attitudes have generally adopted existing instruments, after initial program design but while still in a program's formative stages, to measure variables that corresponded poorly with the program's actual expectations.

6. Coordinating Design With Similar Programs. Nearly all visited programs made deliberate efforts, when offering courses at more than one level, to weave common threads thoughout their offerings so that one course would logically and systematically progress to the next. Most POST-certified curricula include building blocks to be expanded in later courses, thus ensuring both differentiation and compatibility among programs. The Kentucky POST programs represent the most elaborate effort at coordination because both supervisory and management sequences include three substantively related courses and a series of workshops, intended for attendance annually. These are capped by a Command Decisions Course for executives. Most organizations began with a single program and found either that its graduates wanted to return for a higher-level course or that new applicants required lower-level courses and, thus, initiated additional courses to meet demand. This diversification in response to demand, especially in independent programs, fosters gradual differentiation among programs.

7. Ensuring Contractor's Design Meets Goals. Resource constraints often dictate that organizations or individuals outside a jurisdiction's training agency be contracted for training delivery. Their selection is rarely based on clear definition of program objectives, however. The process of contracting for services often moves directly from the administrative decision to "go outside" to the solicitation of proposals, and, in the words of one POST training coordinator, "the proposal that is bought is taught." The contracting agency often presumes that goal clarification and identification of program objectives are the contractor's exclusive burden. Training vendors, however, vary in the amount of energy and time they are willing to invest in helping the contracting agency clarify program goals and objectives after a contract has been signed. The Northwestern Traffic Institute's field programs, for example, are known for working with the contracting agency on-site several weeks before delivery to determine local conditions and clarify objectives. In the typical case, the contracting agency makes minimal efforts at goal clarification and objective setting before purchasing services and has little influence over predetermined course objectives and packaged curricula afterwards.

F. STEP SIX: PRETESTING TRAINEES AND/OR THEIR DEPARTMENTS

The sixth developmental step is to "pretest trainees and/or their departments, to determine pretraining performance levels." This step may be further broken down into the three substeps in Exhibit 9.

Exhibit 9
Pretesting Trainees and/or Their Departments

Substep	Rationale
Whether a program focuses on individual trainees or their departments, pretest trainees to measure relevant aspects of pretraining knowledge, attitudes, and/or skills.	Pretesting trainees provides a comparative basis for evaluations of in-program achievement and post-training performance. It also yields information that staff can use to clarify needs of the immediate trainee audience.
Also regardless of program focus, survey co-workers to determine their opinions of the trainees' needs.	To gain a broader picture of the needs in a particular class (including any perceived discrepancies in need that ought to be addressed), it is useful to get a perspective other than the trainees'. Superiors, peers, and subordinates can provide a useful pretraining measure of attitudes, management styles, and performance deficits of the trainee audience.
Where the program focuses on specific departments, obtain pretraining measures of overall organizational functioning on selected performance indicators.	If changes in aspects of departmental performance or overall functioning are clearly defined and plausibly linked to program activities, relevant pretraining measures are essential to assess the extent and areas of impact.

1. Pretesting Trainees For Knowledge, Attitudes, And/Or Skills. Based on site visits, few programs formally pretest trainee knowledge, either to make comparisons with posttests or to provide instructors with a profile of trainee achievement and informational gaps. Even programs that systematically posttest against terminal performance objectives ordinarily do not pretest. Among the agencies that pretest for knowledge are the Northwestern Traffic Institute, in both its long course and some unit courses, and the FBI National Academy. (In the field offerings of the FBI Academy, however, pretests are administered only when occasionally required by the state or local host agency. One rationale for not routinely pretesting knowledge, expressed by several programs, is that the pretest itself acts as a powerful intervention, focusing trainees more on taking notes for a posttest than on taking active part in the training experience.) For three years in the early 1970s, POLEX pretested trainee attitudes using a standardized personality inventory, to demonstrate the short- and long-term effects of program participation on personality traits related to leadership.

The national survey suggests that many other programs obtain advance information about knowledge and skill levels. Some also ask the trainee his opinion of his management style and management problems. Typically, the program operator obtains this information at the initial class session and uses it to structure the course. Because there is no intention of a posttest, these are not true pretests, so we consider them later in the context of implementation.

2. Pretesting Co-Worker Perceptions Of Training Need. Few programs conduct a pretraining survey of the trainees' superiors, peers, or subordinates that would qualify as a real pretest. Based on the national survey, about one program in five tries to obtain a rating from superiors of the trainees' training need and/or management style. Half that number try to obtain similar information from peers and half again from subordinates. Often, the program obtains this information in applications, not through formal surveys. In most cases, there is no real pretest because there is no intention of a follow-up to determine changes in these perceptions.

Two observed programs conducted formal pretraining surveys without intention of a posttest. The Oakland PD, as part of "prework" to prepare trainees for the program and give instructors advance information about them, required each trainee to complete a self-assessment and also asked him to obtain subordinate, peer, and superior assessments of his management style and training needs. Similarly, in one of the first Police Executive Institute (PEI) sessions, Sherwood surveyed trainees and their subordinates in parallel to tap perceptions of the trainees' management styles. PEI conducted this survey on a one-time basis, as part of a class exercise and to generate needs assessment data for future program development. In contrast to these examples, most cross-cutting departmental surveys of training needs do not focus on the individual trainee. Therefore, they do not have any potential use in identifying the individual's training needs or in later evaluating the usefulness of training for individual trainees.

3. Pretesting Organizational Functioning Or Departmental Indicators. Notably little effort is made to develop baseline data on organizational functioning (e.g., organizational structure measured on a standardized instrument) or key performance indicators (e.g., complaints, sick leave) before training. Based on the national survey, agencies are less likely to obtain pretraining measures on key indicators of departmental performance than any other type of advance information, with the exception of subordinate ratings of training needs. The absence of measurement probably reflects uncertainty about how training ought to impact on the agency and against what criteria to assess management performance and agency productivity.

There may well be more measurement of training's departmental impacts than we found, however. If so, it is most likely that pretraining baseline data are collected in large departments where the chief executive is committed to, or is at least contemplating, changes in the management system. For example, the District of Columbia Police Academy (with FBI assistance) conducted a program to familiarize trainees with MBO and obtain their critical feedback before the department would decide whether to adopt MBO. The FBI instructors pretested trainees using a standardized instrument that measures perceived "blockages" to organizational functioning. This pretesting serves two purposes. First, data for the

three management levels--supervisory, mid-management, and executive--can be aggregated separately, allowing instructors to show trainees how the three levels prioritize blockages to organizational functioning differently. Second, pretesting measures of organizational functioning may be compared to later measures: first, after management ranks have been trained in and have experimented with MBO in their own domain, to measure the effects of training and, second, after the implementation decision has been made and followed out, to measure the effects of MBO on departmental functioning.

The decision not to pretest departments is probably appropriate in most situations. This is so for two reasons. First, police managers are generally not selected for training according to individual need. Second, their participation in disparate programs, having varied emphases and endorsing divergent management systems, does not plausibly produce systematic results.

G. STEP SEVEN: IMPLEMENTING THE PROGRAM DESIGN

The seventh developmental step is to "conduct the program in a manner that corresponds with program design and addresses actual trainee needs." This step may be further broken down into the seven substeps in Exhibit 10.

Exhibit 10
Implementing the Program Design

Substep	Rationale
Control trainee selection to ensure that enrollment standards are maintained.	A program cannot achieve its objectives if it fails to select as trainees those who need the training it offers and are capable of attaining anticipated performance levels.
Control the hiring and retention of trainers to ensure that relevant standards are maintained.	By hiring trainers in accord with relevant standards (those that recognize both individual capability and capacity to fit into the program), trainee satisfaction and learning are maximized.
Coordinate trainer activities to ensure initial acceptance of program goals, objectives, and instructional processes and to maintain this commitment.	When trainers first join program staff and later as they continue to serve on it, some formal coordination is generally needed to secure and maintain commitment to a shared set of expectations.

Substep	Rationale
Implement a formal auditing procedure to monitor trainer compliance with program content and process requirements against objective criteria.	Even where trainers are hired according to strict standards, it is generally appropriate to observe the instructional process periodically to ensure continued compliance with program design.
Maintain rewards and sanctions for successful and unsuccessful completion of training.	Effective use of rewards and sanctions promotes in-program participation and learning, post-program transfer of training, and future interest in attending other training sessions. Their ineffective use reduces credibility and detracts from other efforts to achieve program goals.
Obtain advance information on the backgrounds, perceived training needs, and current management problems of trainees, and provide this information to trainers prior to the commencement of the training session.	Advance information can show how the training needs of the immediate class differ from those of the intended audience. It can, thus, provide the basis for adjusting or refocusing planned program activities.
Ensure that trainers respond to the actual needs of each training class by adapting program content and methods consistent with overall goals and objectives.	It is a more effective use of program resources to accommodate training activities and expected training outcomes to the actual needs of the immediate trainee audience than to maintain the original program design inflexibly.

1. Maintaining Trainee Selection Standards. From one program to another, ability to maintain selection criteria varies widely, as does interest in doing so. Control over trainee selection appears strongest when two selection models dominate: pre-service and other rank-related selection and status selection. Control is less consistent when career progression and departmental discretion models of selection predominate. Programs that use the latter two models are likely to become less stringent in control over selection under three circumstances:

- When in formative stages, in the attempt to develop a clientele
- When faced with declining demand, in the effort to "hang on"
- When confronted with budget justifications that require inflating the population of program graduates

The opportunity to exercise control over selection is lowest for "roadshows," or field programs, conducted by academic institutions, professional associations, and the FBI. Because they usually do not arrive on-site until training delivery, field programs depend on the local sponsoring agency to select trainees in accord with stated criteria. In practice, the local agency often assembles a class much larger, more heterogeneous, and lower in average rank and responsibility than promised. Occasionally, the opposite happens and the class is smaller, more homogeneous, and higher in average rank and responsibility than promised and considered desirable. This low control over selection standards is found in all programs that place the definition and enforcement of selection standards in the hands of individual departments or other sponsoring agencies. The deliverers tend to lose control because the local agencies have different purposes for sending officers to training.

2. Maintaining Trainer Employment Standards. Standards for hiring trainers are so varied and allow such latitude for weighing alternate credentials that it is often hard to tell what standards are being maintained. Few instructors in the programs that we observed lacked the minimum teaching credentials required by the programs. Some interesting patterns emerged, however, from on-site observations and the national survey. First, program operators make little effort to determine if new hires support the program's general philosophy of management training. Second, programs claim to base hiring decisions on three considerations: subject matter familiarity, training capability, and trainer-trainee rapport. In practice, most program operators seem content if applicants convincingly meet their demands on two of the three, expecting that "the rest will come out in the wash" through surveys of trainee satisfaction. Third, programs do not base their hiring decision on subject matter familiarity to the extent that one might expect from the high importance attributed on the national survey to "experience directly related to the subject areas." More importance seems to be placed on trainer-trainee rapport than other considerations. Most of the programs that we observed primarily hired instructors with law enforcement experience. They nominally represented those with business or educational backgrounds on their staffs. When a program lacked a unifying theme and clear objectives, there was a tendency to to hire more on the basis of local prominence and "drawing power" than on criteria related to content familiarity or teaching ability. Fourth, based again on the national survey, it appears that most programs put a premium on experience directly related to the subject areas if it was obtained in law enforcement and if a trainer candidate already has prior experience as a trainer. There is a strong tendency toward inbreeding and the rejection of "new blood." These patterns suggest that many program operators do not have a clear picture of the purposes for their program and of how hiring decisions should support these purposes.

3. Coordinating Trainer Activities. On-site observations suggested that most programs spend minimal energy trying to ensure a trainer's initial acceptance of its goals, objectives, and instructional processes and to maintain this commitment. It seemed that, except in small programs, rarely are faculty convened to become more familiar with each other's presentations, identify common themes, and define an effective course progression. It seemed that, at best, coordination is typically performed by one individual, who observes each new instructor

to get a sense for his rapport with the class and to identify unnecessary duplication in content.

Some observed programs did substantially more than this. At POLEX, for example, all new instructors are observed and/or videotaped. The program director sits in on part of each instructor's block of time. Periodic faculty meetings are held to maintain congruence of management philosophy. The size of the faculty is kept relatively small, streamlined from its original 17 down to 5.

The national survey forces us to question our initial conclusion that trainer activities receive minimal coordination. When asked "To what extent are certain means used to coordinate trainer activities and ensure compliance with program curricula?," survey respondents gave a moderate or high rating to each response. They rated three means as highly used; in declining order of importance, they are: observation of each new instructor in class, periodic spot observation of all instructors in class, and periodic review of lesson plans. Again in declining order of importance, they rated these two as moderately used: periodic interviews by the program director with each faculty member and periodic faculty meetings.

How can the difference between on-site observations and the mail survey be explained? The differences may not be as great as they seem. First, extent of use may be exaggerated in the survey because all the alternatives are desirable and respondents wanted to "look good." Second, although the magnitude of reported use is high, the relative frequency of the several means of coordination were in line with on-site observations. For example, the most used means of coordination was observation of new instructors in class. The least used means was periodic faculty meetings. Third, the responsibility for these activities is still likely to fall on one individual. Fourth, the likely focus of these activities is still on reducing unnecessary duplication in content and maintaining trainer-trainee rapport rather than ensuring acceptance of and commitment to the program's overall purposes.

4. Auditing Trainer Compliance With Curricula. Spot observations of individual trainers seem to be performed in most settings. They are done on the initiative of the sponsoring agency or, in some cases, an oversight agency, such as a POST Council. Periodic audits, usually short in duration and focused on the discrete module being presented, can examine physical arrangements and other logistics, content and methods, personal presence of instructor, general atmosphere, nature of trainer-trainee interaction, and so on. However, even when such audits are conducted by an oversight agency, the auditor's assessment is rarely predicated on a clearly defined set of evaluative criteria or systematic observation procedures. The absence of standardization in auditing procedures and evaluative criteria can become significant when several auditors operate in a single jurisdiction. This ties in directly to the point made above, that efforts to coordinate trainers and ensure their commitment tend to be haphazard and focus more on procedure than on the program's purposes.

5. Maintaining Rewards And Sanctions For Training. The effective use of rewards and sanctions for completion of training can be examined from two

perspectives: Does the system promote in-program learning? Does the system encourage post-program utilization of what is learned?

To see if the system promotes in-program learning, we first have to identify the rewards and sanctions that programs potentially control. Nearly all programs directly control issuance of a certificate of completion. A smaller number directly influence academic credit, pay incentives, management certification, and other incentives. To obtain a certificate of completion in nearly any program, a trainee is obliged to attend a specified percentage of classes (generally about 90 percent), although attendance taking is often inconsistent.

Based on the national survey, two programs out of three require that trainees "participate actively in all attended sessions." Three out of five require that trainees pass one or more exams. One in three requires trainees to make a class presentation, complete a group project, and/or display certain desired behaviors to the satisfaction of the instructor. One in four requires that trainees take one or more ungraded exams. One in five requires trainees to write a paper and/or submit a notebook (record of class notes). To obtain academic credit, the trainee may have to fulfill other requirements. In some programs, he may have to attend a specified percentage of classes or pass one or more exams, although he would not have to obtain a certificate of completion. In most cases, based on the survey, a trainee simply has to write a paper and/or pay an additional fee. Most programs report they have experienced no problems in maintaining attendance and that few trainees fail to complete a program once they start. One program, for example, reported only two non-completions in over a decade. In some, it was clear that the cut-off grade on the final exam was adjusted to allow all to pass.

Do programs encourage a high level of performance? One may argue that they do not because a trainee must merely be present in class, perhaps take an exam or two, and maybe complete several other ungraded tasks to obtain a certificate of completion. Except where academic credit is given, trainees are as likely as not to obtain a grade that indicates differentiation in performance. There is a strong counter to the argument that programs fail to encourage high performance. The wide mix of academic backgrounds in a typical class could not only make a more competitive situation unfair but also could discourage officers who have long avoided competitive academic situations from attending at all. A competitive situation could shift the focus of attendees from taking active part to preparing for the examination. This could deter those with different backgrounds from interacting freely and obstruct development of a class network.

The second question, about the use of rewards to encourage post-program utilization, was articulated most clearly by the director of a POST Council. He argued that, "We need to look at the issue of whether people are going to use knowledge or just plaster walls with certificates." On the assumption that programs currently overcertify and fail to see feedback about utilization as an integral component of the training process, he argued that, "People should first know what expectations are placed on them. At the end of training, no credential should be given. Then, six months after training, the individual must demonstrate how training has been used. Only then should training's utilization result in a certificate award." Virtually no programs require evidence of implementation

efforts, as here suggested. Again, there is a strong counter to this basic argument. Expectations imposed on an individual trainee are often difficult to identify or are inconsistent. Much training focuses on shaping effective managers through long-term career development and a parallel incremental build-up of the organizational capability for change rather than on immediate implementation. In this light, it might be more appropriate to require evidence of implementation efforts than evidence of successful implementation.

6. Obtaining Advance Information About Trainees. Many programs try to obtain advance information about trainees. The types of information collected and means for their collection vary, however. When asked on the national survey what types of advance information they collected about trainees or their departments either before or at the start of training, as many programs as not reported the collection of general biographical information. This covered such items as law enforcement experience, educational achievement, and prior related training. One in three collects information on pretraining knowledge or skill levels (sometimes as part of a pretest). One in four obtains the trainees' perceptions of their own training needs and/or management style or has trainees describe their key management problems. One in five obtains superior ratings of the trainees' training needs and/or management style. A like number obtains objectives set for trainees by superiors as to what to gain from training. One in six obtains trainees' self-selected objectives for what to gain from training. About one in ten collects other information, such as pretraining measures of trainees' on-the-job behavior, trainees' personal and career goals, pretraining attitudinal or personality measures, or peer ratings of trainees' training needs and/or management style.

The means for collection of the information vary. Based on the national survey, one program in two obtains it through applications. One in three obtains advance information at the initial meeting of trainees with program staff. Other means are used one time in six. These include individualized testing, review of departmental performance data, direct observation of performance on the job, and formal pretraining surveys. One program in ten collects advance information through assigned prework. To compensate for the lack of true advance information, many instructors conduct exercises early in a course or module through which trainees share their management problems, personal course objectives, and training needs.

"Prework" is a recent innovation in the collection of pretraining information about trainees. It has three purposes: (1) to motivate the trainee to think about his training needs, (2) to help attune key co-workers to the trainee's training opportunity, and (3) to provide program operators with information about the trainee and his agency. PEI has gone as far with prework as most agencies have. It simply asks trainees to describe their management problem in the program application. The Oakland PD, in contrast, calls for a battery of activities in prework. Not only do trainees complete a sequence of readings, but they also fill out a self-assessment of management styles and training needs, a survey of management problems, and a time use study. In addition, trainees are expected to ask subordinates, peers, and superiors for an assessment of trainee needs and management style. Prework can, thus, provide program operators with more

than enough information to accommodate the planned program to the actual needs of trainees.

7. Ensuring That Trainers Accommodate Actual Trainee Needs. Where detailed advance information about trainees and their needs is available, it is used to seed classes and discussion groups, to promote and structure interaction, to focus instruction, and so forth. Based on the national survey, advance information is primarily used in two ways: to adjust the overall level of presentation and to add supplementary exercises and materials that focus on the overall immediate needs of a class. Advance information is much less likely to be used to provide a comparative basis for evaluation, to select trainees and assign them appropriate sessions, to give individual attention to trainees, and to provide a framework for trainee development of personal action plans.

Whether trainers can systematically accommodate actual needs of the immediate audience depends on answers to five questions:

- How familiar are the instructors with their material?
- How extensively do instructors use experiential methods?
- How early in a course, or how far in advance, can instructors obtain information on trainees' backgrounds and particular needs?
- How much discretion do instructors have to switch materials or modify objectives?
- Are training materials available other than those originally planned for use?

Trainers from observed programs attested to the significance of these five factors. For example, the lead instructor in one program stressed the importance of experiential exercises, saying they "allow each workshop to be adjusted to the emphasis needed by the immediate participants." Instructors in a program extensively involved in field training said that they regularly brought along an alternative course outline and materials, "in case the class we find when we arrive is too high or too low for what we planned." And, as a general rule, if trainers encouraged and successfully elicited active trainee participation, the trainees almost forced them to adapt course content to the class's academic preparation and functional responsibilities.

H. STEP EIGHT: PERFORMING IN-PROGRAM EVALUATIONS

The eighth developmental step is to "evaluate the in-program effects of training on participants." This step may be further broken down into the two steps in Exhibit 11.

Exhibit 11
Performing In-Program Evaluations

Substep	Rationale
Assess trainee satisfaction with the program's logistics, contents, manner of instruction, and projected usefulness.	Positive trainee reaction throughout a program is generally a precondition for learning and behavioral change. Its measurement reflects "consumer satisfaction." Although of restricted validity, reaction surveys are easily implemented, are seen as a *sine qua non* in maintaining accountability, readily complement other in-program outcome measures, and often provide useful information about details of program implementation.
Utilizing established success criteria, evaluate in-program effects of training, such as increases in knowledge, improvement in skills, and changes in attitudes.	By measuring the in-program effects of training on knowledge, skills, and attitudes against established criteria, one secures direct evidence of program success relative to objectives that are under the control of program staff.

1. Assessing Trainee Reaction To Training. The much-maligned reaction survey, course critique, or "happiness evaluation" comes closer than other measures to being the universally accepted barometer of training's success. This does not mean reaction surveys measure everything that is attributed to them, only that nearly all programs conduct one. There are exceptions; field programs, including the FBI Academy's field offerings, tend to be critiqued only at the sponsoring agency's request.

Described by one POST administrator as "an exercise in going across the page and filling in a continuum of smiling and frowning faces," course critiques actually vary greatly in emphases, complexity, and the purposes for which they are done. Among the observed programs, some were simple and some complex. In one program, it consisted of a brief group discussion moderated by the program coordinator at the course's end. In another, it consisted of a five-question, close-ended, multiple choice survey given when the course was over. At the other extreme, the complex critiques were phased in several components. In one program, on the next to the last day, the lead instructors conducted a reaction survey asking for detailed responses about each instructor and other training needs. On the last morning, they discussed these written critiques in a class session. The last afternoon, the academy director guided a three-hour oral class critique of the quality of instruction and the implications of class content for departmental decision-making. In another program, after each instructor completed his

block of classes, the director asked trainees to rate each lecture separately in terms of the quality of content and teaching ability. In addition, after each exercise, the director requested a short written critique. This critique asked if the exercise effectively supported a lecture, was germane to management needs, and was worth repeating. At the program's conclusion, the director conducted a group oral critique to obtain reactions to the overall program, based on broader perspectives than allowed by the critiques distributed during the program.

The emphasis found in a critique appeared to depend on whether it was being conducted mainly to assess logistics or instruction. Many critiques that dealt with content, instructional methods, and projected usefulness did so in a global manner. Resulting information could be of little use in improving the program's delivery, although it might well be useful in budget justifications. The thrust of many logistics-oriented critiques was not on understanding the ways in which the course's content and instructional methods might be altered to increase usefulness. Their thrust was ensuring general satisfaction with administrative arrangements and amenities of the setting. As a result, they place heavy emphasis on matters such as housing arrangements, palatability of refreshments and meals, air temperature and seating design in classrooms, hardness of classroom seats, and so on. Programs that conduct critiques for a balanced combination of administrative and instructional purposes also try to tap detailed and constructive trainee reactions to content, manner of presentation, and projected usefulness on the job.

2. Measuring In-Program Outcomes. Based on direct on-site observations and the national survey, there are only two commonly used methods for the measurement of in-program outcomes. These are: a written examination to test changes in knowledge or skill and instructors' structured observations of changed trainee behavior.

Slightly more than one program in two conducts a written examination to test changes in knowledge or skill. Most do so without a pretraining knowledge measure for comparison to show what a participant actually gained in training. Generally, programs with terminal performance objectives at least partially base their exams on them. This is called "criterion-referenced testing." Those without terminal performance objectives are marginally more likely not to test than to test. If they do give a written exam, it is most likely constructed of questions submitted by individual instructors for their own blocks of instruction.

Structured observation of behavior changes is used in about two programs in five. Instructors typically observe changes in trainee behavior in structured group experiences or other simulation exercises. Programs also use three other methods to measure in-program outcomes. About one in ten reported that trainees develop an action plan based on program contents. A like number reported that trainees assess each other's performance. One in 20 said that they measure attitude change on a paper-and-pencil personality test.

Even programs that questioned the appropriateness of a formal knowledge measure (for reasons in the discussion of rewards above) carefully scrutinized changes in interaction patterns and provided trainees with informal feedback. In one of the programs that we observed, for example, the instructor did no

formal testing. He closely observed changes in structured group experiences, however. He gave trainee-graded pre- and posttests, or posttests alone, to sensitize trainees to new concepts and demonstrate to them what they had learned. This program and many others considered a formal and public measure of outcomes to be unimportant and even counterproductive.

I. STEP NINE: PERFORMING FOLLOW-UP EVALUATIONS

The ninth developmental step is to "evaluate the effects of training outcomes on the work setting in terms of trainee performance and/or organizational functioning." This step may be further broken down into the two steps in Exhibit 12.

Exhibit 12
Performing Follow-Up Evaluations

Substep	Rationale
Follow up on trainees after they return to their jobs by measuring how trainees and their co-workers view the projected usefulness and actual utilization of training.	Follow-up surveys of trainees and their co-workers are generally the most feasible means of evaluating the transfer of training to the work place. They can be especially informative when compared with preprogram surveys focused on individual trainee needs.
Use established success criteria to measure empirically the post-program utilization of training and its impact on trainee performance and/or departmental functioning.	Empirical measures of training's effects on the work setting, difficult as they are to obtain, can be useful when certain conditions are met. Program expectations must be well defined and plausibly related to program activities. It must be likely that resulting information will be used by program operators or oversight agencies. The costs of measurement must be reasonable in light of the ways information will be used and the resources allocated to the program itself.

1. Measuring Follow-Up Perceptions Of Utilization. Most programs rely on informal and often unsolicited feedback from individuals who have applied training, or from their superiors, to provide illustrations of the practical applicability of their offerings. Based on the national survey, many programs also conduct two types of follow-up evaluations after trainees have returned to their jobs and have had a reasonable chance to use what they learned. These are: the trainee's self-assessment of training utilization and assessments of training utilization from superiors, peers, or subordinates. An equal number of programs,

about one in three, reported the use of these two approaches. Among the agencies that we observed, several conducted utilization surveys for some of their courses. These included the PEI, the National Sheriff's Institute, the FBI National Academy, Northwestern University's Traffic Institute, and the South Dakota POST Council. Although we cannot be certain, it seems that many of these follow-up surveys are viewed more appropriately as "rebound evaluations" than as utilization surveys. This is because many do not focus on actual use of what trainees learned as much as they reiterate the course critique--as if to ask, "Now that you are back on the job, do you still feel as good about the course as you did before?"

Based on the national survey, other follow-up evaluation approaches are also used. About one program in five had trainees reassess their training needs, obtained similar information from co-workers, or had trainees assess changes in their management styles. One in ten had co-workers assess changes in trainees' management styles, and a like number obtained the trainee's self-assessment of progress in implementation of a personal action plan. About one in 20 had co-workers assess progress toward implementation of the trainee's action plan.

2. Measuring Transfer Of Training And Agency Impact Empirically. The foregoing sections should offer little reason to expect that programs have empirically measured the transfer of training or agency impact. Based on the national survey, about one program in ten relies on measures of overall organizational productivity to evaluate training. About one in 20 uses management audits or structured observation of actual on-the-job behavior. Almost no programs use cost-benefit analysis, paper-and-pencil assessments of changes in organizational functioning, structured observation of simulated job situations, retesting to determine retention of course contents, and measures of attitudinal change on a paper-and-pencil personality test. In other words, most programs have performed no empirical evaluation of training's effects on job behavior and the work setting.

In addition to external factors that inhibit evaluations of programs, there are internal factors that make evaluation difficult and generally not worth its cost. These include: lack of information about the training needs of the target population, obscurity in in-program objectives, unclear expectations about how in-program outcomes transfer to the work setting, selection of trainees on the basis of criteria other than demonstrated need, lack of effort spent in clarifying for trainees what they are expected to learn in training and carry back to the job, the dispersal of trainees to multiple programs with varied orientations and often with divergent notions of desired training impact, and the near total absence of pretesting. In Chapter Five, we will consider the implications of these factors and others for program evaluability.

J. STEP TEN: USING EVALUATION RESULTS

The tenth developmental step is to "use evaluation results in subsequent program development and revision." This step may be further broken down into the six substeps in Exhibit 13.

Exhibit 13
Using Evaluation Results

Substep	Rationale
Share evaluation results with those who provided inputs to goal setting in order to focus their succeeding goal inputs.	Those who contributed to the initial goal-setting process deserve ready access to evaluation results. This will maintain their commitment and focus their subsequent inputs upon the evidence of effectiveness.
Use evaluation results to refine the goal-setting, needs assessment, and objective-setting processes.	By documenting the extent to which the the program attains its goals and objectives and meets trainee needs, evaluation results can help refine the processes by which goals and objectives are set and trainee needs determined.
Use evaluation results to identify and modify particular goals and objectives that do not correspond with documented trainee needs.	By documenting the extent to which the program meets trainee needs, evaluation results can identify particular goals and objectives that should be changed because they do not reflect trainee needs.
Refine program components, including training staff and curriculum elements, based on what evaluations reveal about their effectiveness.	Program design and delivery can be improved based on evaluation indications about the extent to which program design was implemented and about the effects of implementation on attainment of objectives and overall goals.
Use evaluation results to identify and reduce the influence of factors external to the program that impede its systematic development.	Evaluations can sometimes reveal exogenous factors that impede program development and limit attainment of program objectives. Their identification and understanding can lead to strategies for reducing their potency.
Use evaluation results to justify educational efforts aimed at reconciling discrepancies over the desirable outcomes of training.	Where evaluation results help clarify training needs, they can be used as a springboard for educational efforts aimed at reconciling discrepant views of training needs.

1. Sharing Evaluation Results With Contributors To Goal Setting. It is not clear how extensively and in what forms programs share evaluation results, including course critiques, with those who originally contributed to the goal-setting process. For most programs, the original goal-setting process is rather closed. It does not invite contributions from diverse criminal justice and community interests. However, some programs receive formal input from their advisory boards and share evaluation results with them. Programs that serve a single municipality or county share evaluation results with departmental officials. The District of Columbia Police Academy's director, for example, provided a summary of the oral course critique to the department's executives, to help the department make a more effective decision about implementation of changes in its management system and philosophy. In most programs, however, if evaluation results are disseminated at all, they appear to be shared informally and selectively.

2. Using Evaluation Results To Refine The Objective-Setting Process. Based on site visits and the national survey, it is apparent that programs do make changes in their goal-setting, needs assessment, and objective-setting processes on the basis of evaluations. The national survey shows that evaluations are used in this way to a moderate extent, which means that programs still make extensive use of evaluations only to modify specific course components. To a moderate extent, they use evaluations for other purposes. They are most likely to make moderate use of evaluations by changing specific goals and objectives. With lower frequency and to a lesser extent, they use evaluations to refine the entire process of assessing needs and setting objectives.

On-site observations help explain how this refinement occurs. Typically, reaction survey results showed that training needs had been inadequately reflected in the program's original goals and objectives. This evidence suggested the need for procedures to obtain systematic and broader input into these processes. To do so, programs implemented formal periodic needs surveys. One program even started using an executive training course as a sounding board for statewide executive input.

The national survey supports the observation that the predominant change in the needs assessment and objective-setting processes involves efforts to obtain more systematic input from user groups. To a relatively low extent, they have used evaluations specifically to "change the nature or extent of contact with user groups prior to training."

3. Using Evaluations To Modify Goals And Objectives. As we explained above, programs use evaluations most extensively to modify specific course components. They use evaluations to a high extent in no other ways. They make the greatest moderate use of evaluations by modifying specific goals and objectives. From what we observed on-site, it seems that most program operators identify new objectives and modify inappropriate ones indirectly. When they shift topical focus, course objectives change along with them. Occasionally, successive shifts in topical focus have the cumulative effect of producing a unifying theme and of, thus, clarifying goals. Based on the national survey, program operators also use evaluations to a moderate extent in the development or expansion of training options at other levels of management.

4. Using Evaluation Results To Refine Program Components. Program operators use evaluations extensively to "doctor" their offerings. We do not mean to imply any misrepresentation by the word "doctor," only that the focus of these changes reflects an overall "band-aid approach" to program development.

Based on direct observations and the national survey, evaluations are used most extensively in five ways. They are used to alter instructional techniques, expand or increase the use of particular instructional personnel, revise and update the course description, modify the order or sequencing of course modules, or eliminate unpopular courses and change topical emphases. To a moderate extent, evaluations are used to alter logistics (housing arrangments, meals, class seating) and to change trainer hiring standards. To a small extent, evaluations are used to implement or refine auditing procedures or to increase emphasis on transfer and implementation techniques. They are hardly ever used to change the course name or to change trainee selection standards.

Many programs undergo change regularly. The program operators change the contents of existing modules, including exercises and materials. Based on popularity and demand, they add certain modules or expand their scope, while they drop or reduce the scope of others. Depending on how trainees rated instructor performance, the operators increase or decrease the use of particular personnel. Gradually, topical emphases shift. Staff composition also changes gradually as unpopular instructors are weeded out and even de-certified. Administrative arrangements and logistics, meanwhile, might undergo almost constant revision.

Some program operators also use exams as an immediate check on the quality of instruction. They identify frequent errors on the exams that, they assume, reflect areas in which instruction was weak. If time permits, they go over the items missed by a large percentage of trainees in the immediate session. If not, they at least try to bolster related modules in subsequent presentations.

Evaluations rarely lead to major program revision. However, a follow-up reaction and utilization survey conducted for the New England Institute of Law Enforcement Management, located at Babson College, occasioned significant changes in staff composition. The survey, directed to several hundred graduates, showed that instructors without law enforcement background were rated not only as having better grasp of content and better instructional styles but also as being more practical than those with law enforcement experience. As a result, the program operators made major changes in staff composition. They dropped most of the instructors with law enforcement experience. This case represents an exception to the rule, however. The lack of information about training's effects on the work setting has generally discouraged major shifts in approach to police management training.

5. Using Evaluation Results To Control External Influences. The national survey showed that program operators and managers use evaluations to a moderate extent to "reduce the inhibiting effects of forces external to the program on program development and delivery." On-site observations help to understand how this occurs. In one program, reaction surveys showed that agency officials expected one type of program and trainees saw the need for another. Trainers shifted their topical emphases toward more innovative content. Course critiques then showed that trainees anticipated a high level of resistance to their efforts

to implement what they had learned in training, a management system incompatible with current departmental philosophy. To reduce the bases for resistance, the trainers held a seminar for the agency's executives. Here they asked the central question: Should the agency shift its management philosophy, and, if so, should the trainers provide related training to all departmental personnel? The use of evaluations to control external influence is generally difficult to document because it happens informally.

6. Using Evaluations To Justify Educational Efforts. According to the national survey, program operators and managers also use evaluations to a moderate extent to "start educational efforts to reconcile discrepant views of training needs." This use is also poorly documented, and with good reason. When someone starts an educational effort to increase compatibility among the expectations of several groups, he stands to gain little by underscoring these differences. The differences initially come to light only informally because prevalent needs assessment and evaluation approaches do not reveal disparate perceptions of training need and utilization. The process of influencing views of training need and creating a consensus is also generally informal. By way of exception, the executive seminar cited above served as a public forum for exploring and reconciling divergent expectations.

* * * *

So, how closely do police management training programs follow the industrial model of program development? Our first reaction has to be that they do so, at best, inconsistently. We can see this by reviewing our appraisal at selective points. Program developers and operators set goals largely without substantial input from user groups. They do little or no formal needs assessment. They typically skip over the identification of performance deficiencies and often pass off topical interest surveys as needs assessments. They do not set consistently clear objectives. They generally do not identify criteria that indicate, in measurable terms, the areas in which change is desired and the intended extent of change. They usually cannot pinpoint how they want in-program outcomes to affect job behavior. They rarely specify the learning principles that underlie instructional methods. They often fall back on an established curriculum long after its use has become counterproductive. They rarely provide a mechanism to help trainees and their superiors come to agreement about the individualized purposes for participation in training. They hire trainers based more on anticipated trainer-trainee rapport than on subject matter familiarity. They tend to exert little control over the composition of a class. They make minimal efforts to coordinate trainer activities. They distribute rewards for training almost indiscriminately. They obtain too little advance information about trainees for it to be useful in targeting content or measuring outcomes. They conduct few evaluations other than course critiques. They use evaluation results to tinker with program components but rarely to make needed major revisions.

We can also answer this question in another way if we step back and take a broader view. The program development practices described above are not unreasoned. The process is deliberate, phased, and rather systematic. It corresponds inconsistently from point to point, however, with the prescribed

steps of the industrial model. This description of practice partially verifies the claim by one program administrator that, "Programs are not rationally designed. Instead, they evolve--are gradually shaped by what is needed." But only partially.

It is just as important to ask, "Why are there differences?" The differences and similarities between the industrial model and current practice stem partly from factors internal to the program. These include divergent assumptions about what is feasible in program development and needed in management training. They also stem from factors external to and outside the control of the program. These include funds, training mandates, and the availability of trainees and instructional personnel. The external factors that reduce one's control over the program development process are the subject of Chapter Three.

Three: Factors Affecting Program Development

How systematic can program development reasonably hope to be? Are there factors external to programs and largely beyond the program developer's control that affect how closely developmental practices can correspond to any chosen system?

To understand the factors that can impede or facilitate systematic program development, we might start by examining the simplest organizational training arrangement, the departmental trainer who has sole responsibility for training design, delivery, and evaluation. In this situation, we find that development is affected by:

- The trainer's general knowledge of training standards and management principles, specific knowledge and skill in implementing training techniques, and willingness and fortitude to abandon the strong customs and precedents in police training
- The trainer's particular views of what is feasible and necessary in program development and needed and effective in police management training
- The funding, legal, departmental, and community constraints that limit discretion
- The supply of already existing program resources
- The perceived or actual scarcity of technical resources, including performance measures and research designs, especially for use in needs assessment and evaluation
- The availability of trainees who fit the program's target audience
- The immediate needs and expectations of the trainees who appear for the program
- The receptiveness of trainees to evaluation techniques centered on their performance

In the typical training situation, however, not one but several individuals or organizations take part in the developmental process. These include a POST council, an SPA, local colleges and universities, national or regional professional or training organizations, management consulting firms, and others. To the list above, therefore, these factors must be added:

- Perceived or actual scarcity of personnel resources to take part in needs assessment, design, delivery, and evaluation
- Complex inter-organizational arrangements and communication flows
- Competing views--on the part of training program developers, other training personnel, and those in a position to influence resource allocation--about what is necessary and feasible in program development and what is needed and effective in police management training

In addition, regardless of the complexity of the training arrangement, how far the program developer adheres to or departs from a particular developmental system at each step affects the options that will later be open because each step provides input to the next.

The profile we have presented above encompasses factors both internal and external to a program. When the typical program developer tries to explain departures from a particular developmental system, however, he focuses on external factors and looks on internal factors as side effects of the external ones. Developers and operators commonly view their programs as surrounded by external factors that force choices and limit options. They define expectations about "the environment," "force fields," "exogenous factors," or "system influences," whatever they be called, at least as clearly as expectations about how the industrial model should be adapted, what training should accomplish for trainees and their agencies, and how best to accomplish those ends. These external forces can directly or indirectly impede or facilitate the implementation of a particular practice. They can impel the program developer to accept knowingly what he considers "second best." Or they can cause him to make unwitting assumptions about what is feasible, necessary, and effective in development of police management training programs.

In this chapter, we describe the major external factors that affect development of police management training programs and show how these factors can facilitate or impede program development. We group the factors in four parts according to the locus of their impact. The four points of impact are: overall program development; goal identification, needs assessment, and objective setting; design and implementation; and the conduct and use of evaluations. Many factors are not empirically independent but can, instead, be seen as causally related to each other or a third factor or as subsets of less precisely defined variables, as will be evident throughout the chapter. The information we present here can be used to assess the strength of the forces that affect police management training programs generally, to identify the forces that affect a given program's developmental practices, and to assess the feasibility of overcoming the forces that affect program development.

A. FACTORS AFFECTING OVERALL PROGRAM DEVELOPMENT

Five general factors affect programs at each major phase of development. These are: funding, legal requirements, the organizational environment, the community environment, and pre-existing program materials.

1. Funding. This factor defines the basic resource limitations within which a program must operate. Funding exerts obvious influence over the entire range of program developmental activities. It can include resources necessary to pay staff and instructor salaries, rent and operate training facilities, provide food and transportation services, maintain rooming accommodations, provide training materials, and reimburse trainee salaries. We can analyze funding's effects on program development in terms of funding levels, funding continuity, and the efficiency of the funding mechanism.

a. Funding levels. The funding level involves the absolute amount of funds available for program development. It also refers to any restrictions placed on the parts of program development for which funds may be used and to the relative accessibility of funds. The sheer size of a program budget has obvious implications for the resources that can be devoted to program development. Any shortfall in the budget's size, actual or presumed, influences how the program developer will downplay certain activities he regards as "non-essential." Low funding levels are also cited often as the prime reason for offering only one level of management training rather than an integrated succession of courses, especially to justify one-time offerings.

The granting or appropriating agency often will place explicit restrictions on the parts of program development for which funds may be used. Typically, it earmarks funds for certain aspects of delivery, such as staff salaries, transportation costs, and reimbursement of trainee salaries. One POST director commented, "It is not the absence of funds. The SPA is always offering us money. It is the lack of staff and a delivery system." Based on the national survey, funding restrictions have their strongest effects on program delivery and evaluation. They have an appreciably weaker effect on goal setting, needs assessment, objective setting, and curriculum design. The lower perceived effect on these activities, however, may only mirror the lower priority that program developers tend to place on them.

Funds for management training are less accessible than those for basic recruit training and other forms of advanced in-service training in some jurisdictions. In some places, all mandated courses must be conducted or at least scheduled before residual funds may be used for other purposes. Because state law rarely requires management training above the supervisory level, management training is at a disadvantage when it is has to compete with mandated programs. In some places, there are also definitional problems that affect the use of allocated in-service training funds for management training. One state in particular reimburses departments for up to 40 hours of in-service training annually per officer. Whether a course defined as management training could qualify as a mode of in-service training had not been resolved, however, which discouraged and curtailed departmental use of management training. Based on the national survey, these limits on the accessibility of funds have little effect on the typical program but pose real problems for a small number of programs.

b. Funding continuity. The consistency and certainty with which program operators may anticipate support from one year or even one session to the next constitutes funding continuity. It generally influences capability to use deliberation in program development. It specifically affects ability to project long-term training schedules and weave common threads among offerings so that one course can logically and systematically lead to the next.

To increase funding continuity, many state programs have tried to protect themselves from the annual cycles of budget justification and outreach for participants and, thus, to provide a shelter from funding cutbacks. They have used two approaches: passage of legal requirement that a program be offered and shift from dependence on annual appropriations to a penalty assessment fund. Program managers and operators see a legal requirement as a guarantee that their program will be on an equal financial footing with other mandated programs and, hence, be no more susceptible to budget cuts and reduced public expenditures than they are. Managers and operators generally expect the derivation of program support from penalty assessments to place funding beyond the vagaries of the normal appropriations process.

Neither of these strategies has been successful. Almost no state POST councils have obtained a mandate for management training at higher than a supervisory level. A penalty assessment fund, for those states that derive support from this source, has not guaranteed program continuity. Because programs have generally failed to reduce the risks of funding discontinuity, management training programs are susceptible to the normal vicissitudes of the appropriations process, are more vulnerable to cutbacks than are required programs, and probably take more staff time on budget justifications than do required programs. Where the risk of discontinuity is apparently low, due to certain "guarantees," periodic reductions in public expenditures can still threaten continued support. Hiring freezes, for example, directly affect ability to retain appropriate trainers; they indirectly affect program attendance by reducing police agency staffs and creating consequent problems in covering for those who attend training. Funding caps have affected even state training agencies drawing support from a penalty assessment fund by limiting the portion of the fund that could be expended. Based on the national survey, two factors have relatively strong effects on the adequacy of program funding: reduction in state allocations to training and imposition of a cap on expenditures.

How a climate of uncertainty can affect planning and continuity is shown by a program serving a six-state region and threatened with termination of support. The uncertainty of funding bred rumors that the program had been or soon would be terminated. At least two of the six states served by the program started to develop local management training alternatives, partly in anticipation of the program's termination and partly from the desire for cost-effectiveness. Although community support helped save it, the program took nearly a year to reestablish itself fully and to regain its former attendance levels and assurances of financial support.

c. Efficiency of funding mechanisms. The complexity of procedures and the amount of maneuvering required to obtain available training funds constitute the funding mechanism's efficiency. Most program operators tend to equate perfect efficiency with a guarantee of complete support from a single funding source. They expect this guarantee to have the dual effect of providing continuity while minimizing staff time spent in negotiations.

The programs that we observed claimed to illustrate several modes of inefficiency. Two municipal departments, for example, shared a common circumstance: lack of a distinct training budget. Their training directors had to develop a work plan and formally request support from departmental operating funds for

each proposed course as need arose. One training director, noting the impact of this circumstance on capability for long-term planning, likened this process to "asking someone to wave a magic wand and come up with the money." The national survey showed that these two departments are not the only ones in which a lack of a budget line item for training affects the adequacy of funds. Lack of a distinct training budget has little or no effect on the typical agency, however.

Another visited program drew its support from multiple state legislatures and POST councils, each with its own perceptions of training goals and objectives and each with its own procedures for funding application. Because the program contended yearly with six funding agencies, program staff likened its situation to dealing with "six mothers-in-law" just to maintain current funding levels. Based on the national survey, dependence on multiple agencies for financial support has moderate influence on the adequacy of program funds.

Both situations--the lack of a distinct training budget and dependence on multiple funding agencies--can pose problems for continuity and certainly are likely to draw heavily on staff time for negotiations. They can also have advantages for program development, though of different sorts. If the organizational environment permits, the lack of a distinct budget can boost chances that training will be conducted on the basis of immediate demonstrated agency needs rather than to fulfill the functions of an institutionalized program. Dependence on more than one funding agency can provide greater assurance that the program will endure in some form, although this can also make it harder to project the program's size. Therefore, what many programs seem to lose through a funding mechanism's inefficiency, they can gain in other aspects of program development.

2. Legal Requirements. This factor defines the formal requirements with which training activities must comply to obtain and maintain certification and/or operating levels. Embodied in the statutory provisions or administrative rulings of a POST council, or in the enabling laws for individual law enforcement agencies, legal requirements can influence the entire process of program development. They do so to the degree that they either prescribe how a developmental step is to be performed or obviate the necessity of a step altogether. Based on our observations of operating programs, the precise nature and extent of their influence seem to depend upon whether: they were developed following a closely reasoned and systematic model; they are set forth by the oversight agency as the minimum basis for program development but not as the full scope of developmental activities; they are regularly reviewed on a formal basis to assess their current applicability; they include a mechanism for their enforcement. Our on-site observations caused us to question whether the typical program meets any of these criteria; yet, respondents to the national survey seemed convinced that the legal requirements affecting their programs meet the first three criteria and do not fall far short on the last, the inclusion of a mechanism for their enforcement. We cannot reconcile these two views, but the interplay of these criteria will be apparent in the following discussion of requirements for objectives, curricula, audience, trainer credentials, and participant incentives.

a. Objectives. Legal requirements in states that certify management training specify programs' objectives in either of two ways. They do so directly by spelling out terminal performance objectives or indirectly by summarizing

course content module by module. There is great variation among states, however, in whether POSTs explicitly set forth objectives as the minimum basis for ensuring that managers possess a core of uniform information or as an exhaustive statement of intended outcomes. There is similar variation within states in how particular offering institutions view the objectives: as the minimum basis for further development and elaboration to meet local needs or as an exhaustive statement of a course's full scope. The process for derivation of requirements also varies, from empirical statewide goal-setting and needs assessment procedures to the assembly of ostensible experts around a conference table to obtain a consensus of their opinions. When the POST issues requirements, it generally recognizes the need for periodic review, but its commitment to a formal review tends to be forgotten due to the more pressing obligation to update mandated programs. Where a real mechanism for enforcement exists, the expectation of enforcement still varies among offering institutions.

The effects of variation in how POSTs require their programs to address certain objectives were evident in programs that we observed. One POST based objectives on statewide goal-setting and needs assessments and explicitly set them forth as the minimum basis for further program development. Offering institutions often conducted their own needs assessments to develop non-required objectives. Required terminal performance objectives included behavioral success criteria, so complying programs had to shift their instructional techniques toward structured experiences and other simulation exercises that permit demonstration of these behaviors. Another POST derived module-by-module course summaries from expert consensus and pretested them in 1970, with the intent of triannual review. It did not explicably set them forth as minimum requirements. The first review started nine years later. The POST's inability to enforce the requirements coupled with the inapplicability of course content to produce great variation in implementation. Some institutions struggled to follow the rules literally. Others claimed to follow requirements "to the letter" but candidly admitted that the requirements allowed enormous latitude, were not enforced, and had little impact on program operations. A POST administrator in the second state commented on the potentially adverse effects of the state's requirements, noting that, "In 1972, we were considered good. Since then, there have been quantum leaps in development in the field and in participants. But we have not changed. We cannot stand on our feet and say our requirements are job-relevant. And they may be so irrelevant, they are counter-productive." These examples show that requirements can enhance program stability, standardize offerings throughout a state, provide a minimum basis for further program development, and motivate trainers to update their content and instructional methods. They can also restrict program development to the breadth or narrowness of their originator's horizons and guarantee their own eventual program obsolescence.

b. Curricula. Little need be noted about curricula that we have not already suggested about objectives. This is largely because most programs do not clearly differentiate objectives from curricula. One point warrants repetition, however. To the extent that oversight agencies do not scrutinize required instructional methods and training materials periodically for their continuing relevance, they risk mandating methods and materials that are not current and may be less than optimally effective.

c. Audience. Legal requirements often define the audience for which a course is intended, to allow resources to be focused on the target population's capabilities and needs. Requirements tend to be explicit when selection follows the career progression model. The last chapter mentioned that POSTs emphasize rank, responsibility, and completion of prior installments of a defined career sequence in this model. They also use other selection criteria, including prior education, other prior training, time in grade, and time elapsed since last participation in the sequence. From state to state, however, there is variation in which criteria are used, how they are prioritized, and how freely priority criteria are waived to permit a class to be filled.

Typically, a POST seems to set forth criteria as an exhaustive statement of eligibility requirements, and programs view them as such. It is not clear from our research, however, whether most POSTs regularly review criteria for their current applicability or whether they instead weaken enforcement when applicability is strongly questioned. One POST required offering institutions to use rank as the sole determinant of eligibility for participation. In conforming with this requirement, one observed program required extensive "prework" activities, including program readings, to equalize differences among trainees caused by the insensitivity of selection criteria. Another POST specified multiple criteria, but many offering institutions found it difficult to fill classes. This was partly due to two other factors: coverage problems in departments served and intense competition among state-certified programs for the same audience. Faced with apparent exhaustion of demand by the specified audience, these institutions increased outreach activities and encouraged application from officers who did not meet the POST's criteria and for whom the criteria could be waived. In the second case, the POST did not formally review the relationship of criteria to the current demand. Instead, it waived criteria to the extent needed to maintain the active status of certified institutions.

d. Instructor credentials. A POST or other oversight agency generally sets forth instructor credential requirements as minimums, and to these the offering institution often adds its own stipulations, sometimes without differentiating between the two. How regularly and thoroughly the POST or the offering institutions review instructor credential requirements we cannot determine from our research; however, it appears that their review is no more regular than that of certified course objectives and curricula and no more thorough. One POST, for instance, described how it was tightening its requirements by decertifying all instructors who had not taught a specified number of hours in the past year. It did this with little consideration of potential needs for the underutilized instructors or of how massive decertifications of this sort choke off the infusion of new blood.

Credential requirements are often broad enough that programs can interpret them to meet the instructional needs of any given training model. Because this is not always the case, however, credential requirements tend to reinforce the continued existence of a particular training model indirectly, sometimes long after the model itself has officially been abandoned or at least moderated and mixed with other models. This is particularly obvious when trainers must have law enforcement background but the curricula focus on management techniques derived from business and industry.

e. Participant incentives. Many oversight agencies are authorized to grant incentives to participants and their departments. Participants might obtain certificates for course completion, management certification, and pay supplements; departments may receive reimbursement of tuition costs, per diem expenses, and even trainee salaries. These incentives can act as a powerful impetus for officer participation in training and departmental release of officers to attend training. POSTs and other oversight agencies set forth these incentives as "the carrot," or the minimum necessity to lure the typical officer into training. They intend this external incentive to be supplemented by the offering institution's capacity to foster intrinsic incentives. We cannot determine from our research whether POSTs periodically review their incentive programs to determine if the nature or amount of incentives require modification. It is clear, however, that, although incentives promote attendance, their extensive use at least partially shifts the source of motivation from an internal to an external locus. Offering institutions that view these extrinsic incentives as sufficient to maintain motivation and to encourage professional development also reinforce the practice of garnishing certificates without emphasis on utilization.

3. Organizational Environment. This factor encompasses the activities, priorities, personalities, and structure of the larger organizations within which a program develops and the relations between the program developer and the larger organizations. The organizational environment can influence the developmental process at any point by determining the immediate circumstances that surround and shape it. The effects of organizational environment can be analyzed in terms of organizational command structure, organizational training priorities, continuity in senior staff support for a program, accord between the program developer and agencies served in how they perceive training need, and host institution requirements.

a. Organizational command structure. The centralization of decision-making and the fluidity of communications influence program development. They determine how much discretion the program developer has to follow his chosen sequence of program development steps. The autocratic command structure of the traditional police agency has centralized decision-making and restricted the fluidity of communications. It has, thus, often used training as the intermediary between senior officials and the rank and file and curtailed the program developer's discretion to follow a systematic developmental process. The training director in a major municipal department where discretion in program development was highly circumscribed remarked that, "We are too busy running from one fire to another," as identified by senior officials, "to devote time to adequate development."

A traditional command structure, thus, often obviates the need for goal identification. It largely determines program goals and, similarly, forms the program's topical emphases, instructional methods, training staff, target audience, and evaluation approaches. In contrast, in more participative organizations, the management philosophy of senior officials is a major consideration in goal identification, but the program developer tends to have more discretion to develop programs in accord with a chosen developmental sequence while remaining responsive to overall agency goals. The hallmark of a participative organization, in fact, would be the development of program objectives through substantial input from line managers, working from the bottom up.

b. Organizational training priorities. The importance that agencies attribute to the training function, management training in particular, can affect the entire developmental process. Agencies reveal their training priorities in the resources that they allocate to develop their own training capability and in the way that they use outside programs. Specifically, the priority that an agency places on training can be seen in:

- The quality of personnel assigned to manage training activities
- The quality and amount of other personnel and material resources provided to the training manager, relative to the organization's overall resources
- The place of the training function in the organizational structure
- The extent to which training activities are coordinated with other organizational activities and included in the organization's planning calendar
- The degree to which senior officials take active part in training delivery and use training not merely to communicate policy but also to help make organizational decisions
- The extent to which trainees are selected to attend outside training programs based on performance and for purposes of career development, rather than based on other factors and for the prime purpose of boosting morale
- The frequency and extent to which the organization will free personnel to attend outside training programs and provide coverage for them
- The level of tuition funds allotted to attend outside programs relative to funds available

We repeatedly heard POST directors and program operators describe training's role and the trainer's job in upgrading police practice as at or near the bottom of the priorities list in most organizations, as an "afterthought" or a "necessary evil," with minimal draw on departmental resources. Several POST directors noted that, in their experience, a police agency's training director is often selected for this assignment not because he possesses special skills but because he can no longer handle street work. They also noted that the training function is typically managed by a first-line supervisor operating within the Administrative Services Division rather than by a higher-level manager operating closer to the chief executive. Although agencies generally regard basic training as a vital means to screen new recruits, incorporate them into the ranks, and convey to them tangible skills in the protection of life, property, and human rights, they still consider management and other forms of in-service training insignificant compared to other uses of agency resources. The low priority that they place on training can affect not only the entire developmental process but also later organizational receptiveness to new technologies and principles acquired through training.

c. Continuity in senior staff support. Power shifts in an agency can precipitate changes in command orientation and training priorities. Hence, any actual or prospective shift in power can change senior staff support for an actual or planned program and jeopardize its existence and resource allotment. This situation can place the program developer in the position of repeatedly "selling" a program to senior staff, with little basis for long-term planning. Discontinuity in senior staff support can also influence impact evaluations. This is because changes in power and policy readily weaken controls needed to ensure that observed changes in performance can be attributed to training and not to other factors.

d. Agreement in perceptions of training need. How closely the program developer and the agencies that the program serves agree or disagree over training needs can affect the entire developmental process. The results of this relative concord or disparity depend on who is aware of it--the developer or the agencies served.

If the developer recognizes that the agencies for which the program was designed interpret their needs differently, it is likely that he will downplay, rather than underscore, these differences. Thus, he states goals for the public, using broad generalities, and avoids needs assessment and evaluation procedures that are likely to reveal divergent measures of effective performance. For example, one program saw as its mission the adoption of participative practices by agencies in a particular state. At the same time, it recognized that the concepts that its agencies were accustomed to employ in describing management performance would have "made a needs assessment at best irrelevant and at worst threatening." They subsequently sought the confidence and support of served agencies gradually. At first, the state police sent two "spies" to attend consecutive program sessions. Based on their feedback, the state police made a commitment to send all lieutenants for training. During training, the lieutenants noted the need to involve senior personnel, if they were to have realistic chances of using what they learned. In a subsequent three-day informational seminar for senior staff, only two of twenty participants spoke, but, during breaks, many noted the program was "on target." Based on this experience, the program developer concluded that, had they set forth their own perception of training needs too aggressively, they might easily have jeopardized chances for eventual cooperation.

In some cases, the agencies served are more aware of divergent perceptions of training need than are program staff. Where other options are open, agencies in this circumstance might require trainees to attend programs that more closely mirror the agencies' own perceptions of training needs. This reduces trainee availability for the program with unacceptable views on training needs.

e. Host institution requirements. Programs housed in academic organizations often must meet institutional requirements to maintain their activities on site or to secure a credit option. The requirements of a host institution typically affect objective setting and curriculum development, selection procedures, and testing. Based on the national survey, host institutions most frequently modify either curricula or testing procedures. They less frequently modify entrance requirements and procedures. One program, for

instance, was required to add an effective writing course to its curriculum when the host institution observed that trainee writing samples were deficient. A second program required that trainees follow the host university's admission procedures and meet the university's selection standards before they could participate. Many host institutions require that a final examination be given and influence the content, structure, and grading procedures for the exam, as a condition for granting optional academic credit. In each of these situations, the requirements of the host were potentially at odds with what the program developer wanted to do.

4. Community Environment. This factor consists of the interests of the general public and their representatives on political bodies and of the relations between the program developer and these ultimate consumers of training. The community environment can affect program development at any point and is analyzed below in terms of the legal obligation that programs be job related, public demands that programs demonstrate cost-effectiveness, and union demands that its prerogatives be honored.

a. Legal obligation that programs be job related. Programs potentially face two major court challenges, both pertaining to the job-relatedness of training. These involve equal employment opportunity (EEO) and vicarious liability. EEO suits could challenge a program as discriminating in its selection procedures, testing, or contents. Vicarious liability suits could challenge a program as providing inadequate job preparation in minimum required duties and attempt to hold trainers responsible, or vicariously liable, for the effects that inadequate training has upon the community.

Although program staffs cited both court challenges as potential threats and have begun to adjust certain practices in program development accordingly, programs have not faced serious negative consequences as a direct result of either EEO or vicarious liability suits. Several programs in major municipal departments were suspended during EEO litigation, but the suits leading to their suspension challenged the promotional process itself, not the pre-service training activities that accompany promotions. In other words, program delivery was interrupted because promotions were halted, pending the outcome of litigation. Similarly, whereas several public agencies outside of law enforcement have faced vicarious liability suits related to the performance of their line personnel, police management training programs have not, to our knowledge, been so challenged. It also seems unlikely that they will be because the issue of vicarious liability typically arises when line personnel in direct contact with the general public have demonstrated gross inadequacies in performance. Managers come into direct contact with the public in a service role less often. Vicarious liability suits do not pose an immediate threat to management training programs.

In anticipation of court challenges, some programs have focused increasingly on the job-relatedness of course contents. Efforts to develop "legally defensible" or "litigation-resistant" programs have increased the attention paid to needs assessment procedures. This has generated rising interest in task analysis, at least at the basic recruit levels. For a variety of reasons--the lack of mandates for management training, shortage of technical personnel, and the

dearth of measures for assessing the police manager's performance, among others--task analyses have rarely been used to alter management training programs. First-line supervisory programs have been developed on the basis of task analyses more often than programs for higher management levels for three reasons: they are more often mandated, supervisors come into more direct contact with the public, and supervisory tasks can be more readily identified. The threat of court challenges appears to have affected course content as well. Many programs stress the manager's responsibility to operate job-related in-service training activities for line personnel.

b. Public demands that programs demonstrate cost-effectiveness. A climate of austerity in public spending has had two markedly different effects on program development. It has increased emphasis on accountability and changed the type of programs that departments use and that offering institutions conduct.

Many programs have begun at least discussing strategies for improving accountability, if they have not yet taken action on them. These are the three basic strategies discussed:

- Implement pre- and posttests and supervisory ratings of utilization
- Reduce the number of individuals trained and focus on those whose training needs and prospects for later implementation have been documented
- Add an action plan as the final exercise in training, to focus trainees on applying knowledge to their own management problems, and to provide a framework for later follow-up on utilization

Although discussions of these strategies have not produced dramatic changes in practice, it seems likely that the continuing public demand for evidence of program effectiveness will further their adoption.

Interest in cost-effectiveness has triggered public support for development and use of more intrastate regional and local training programs. The public tends to view regional programs as having several advantages over out-of-state residential programs. They decrease the need for the trainees' absence from departments and families for long periods. "They pose fewer problems for family life" and, thus, "do not work such a hardship on the men." They reduce travel and per diem costs, especially where commuting is feasible. They increase the homogeneity of audiences, at least geographically.

In response to these pressures, some major residential programs, including two funded by LEAA through the early seventies, have carried their services off campus, "to the people." These two programs still value a residential concept and espouse a training model much different from that endorsed by many chiefs and POST directors. They recognize, however, that, if they decline to shift partially to a regional format, then agencies will likely purchase regional training services from other programs and might even withdraw financial support for the program's residential component. They would still prefer to stay on campus. When they conduct programs off-campus, control of the training environment shifts to the local host agency or regional academy. This can result in inappropriate environmental features, such as small, crowded classrooms; armchair desks;

lack of blackboards or projection equipment; and intermittent interruption by bells. The purposes for a residential program are not easily adapted to road-shows.

As an alternative to bringing residential programs off campus, some jurisdictions have worked with local colleges and regional academies to develop local management training options. Ironically, even in these cases, they have often imitated the curriculum of a major residential program or substituted needs assessments borrowed from other jurisdictions for a real one.

c. Union demands that its prerogatives be honored. A union's interest in protecting its constituency can affect program development at any point. We did not observe the effects of unions on the developmental process directly, but POST staff and visited programs often reported them. They described five ways that unions can affect program development. They can:

- Control attendance by defining the training day as in excess of eight hours and, hence, justifying overtime pay
- Influence course content and topical emphases
- Promulgate conditions under which training may be offered
- Restructure selection procedures so that no member may be discriminated against on the basis of ability
- Interfere with follow-up evaluations by telling trainees that their program evaluations could be used as personal performance appraisals

The national survey gave no reliable indications about the absolute frequency for these five effects. The absolute frequencies for all five tend to be low. We suspect these understate the reality because they do not fit with what program staff told us on site. This underreporting could occur for two reasons. First, the union influences sound undesirable. Second, the union influences tend to be subtle, and the typical survey respondent may not be aware of them. Still, based on the national survey, unions influence program development in two princial ways. They control attendance by defining the training day as being in excess of eight hours and, hence, justifying overtime pay. They do this especially when the program requires the trainee to stay away from home. The prospect of paying overtime reduces the willingness of chiefs to free officers for training. To a lesser extent, they influence course content and topical emphases. They sometimes insist that a course be given on a "hot topic," despite the absence of demonstrated need. It is not clear from the national survey whether unions extensively affect program development in any other ways.

5. Pre-Existing Program Resources. This factor involves the prior availability of tested program resources and their incorporation into a new program with minimal attention to program development standards. These resources can include goal statements, needs assessments, performance objectives, curricula, success measures, facilities, trainers, and evaluation strategies. Each can be a valuable program input; but, in practice, each can set a ceiling on an aspect of systematic program development. They can exert a powerful influence in

retarding reasonable use of resources. In part, their adoption *in toto*, without regard for relevance to immediate training needs, stems from legal requirements and funding constraints. It also reflects the assumption that "others have done it better, so why reinvent the wheel?" This assumption we label "developmental inertia."

The features of other locally available programs can also influence use of pre-existing resources but in a different way. In many jurisdictions, several programs are available locally, including those contracted directly to agencies. Keen interprogram competition for participants can lead programs to package marketable pre-existing materials attractively rather than go through the steps necessary to ensure that they are addressing real training needs.

The lure of pre-existing program materials is pervasive and seems to affect all programs to varied degrees. The national survey showed that programs fall back on readily available trainers, curricula, and facilities most extensively.

B. FACTORS AFFECTING GOAL FORMATION, NEEDS ASSESSMENT, AND OBJECTIVE SETTING

In addition to the five general factors discussed above that affect overall program development, there are two other factors that influence goal formation, needs assessment, and objective setting. Before we discuss them, it is necessary to put the general factors in perspective. When we observed programs on site, their operators most often cited funding restrictions, legal requirements, and the organizational environment as the factors that affect options during these early stages of program development. The national survey supported and amplified this finding. It asked about factors that "have reduced or even eliminated the need for a more formal process of goal formation, needs assessment, and objective setting.

Based on the national survey, legal requirements and other mandates imposed on the program have reduced or eliminated the need for a more formal process in about two programs in three, quite predictably. Organizational training priorities have had a similar influence over one program in three. The command orientation of the organization that houses the program has deterred a more formal process in one program in five. Actually, if we look at the effects of command orientation more closely, it primarily affects municipal agencies. Fifty percent of municipal agencies have felt the effects of departmental command orientation at this early stage of development, but only one nonmunicipal program in eight has done so.

The more penetrating findings mirror the types of attitudinal variables that were discussed above in the organizational environment section. More than two survey respondents out of three said that program staff and user groups share such a close relationship and understanding on an informal basis that something more formal is unnecessary. About one respondent in two said that the program staff believes that certain needs must be addressed, regardless of whether user groups happen to be conscious of them, so a more formal process would waste effort and could be counterproductive.

Over and above the five general factors that affect overall program development, there are two others, as noted above. The two additional factors are technical resource availability and target population characteristics.

1. Technical Resource Availability. This factor encompasses the availability of personnel, techniques, and other resources needed to conduct early developmental steps. It can be analyzed in terms of availability of relevant concepts and measures, research techniques and designs, adjunct data sources, skilled personnel, and lead time. Based on the national survey, these five factors affect "ability to use a more formal type of needs assessment" to a moderate and roughly equivalent extent. Vastly more extensive an influence than the shortage of technical resources, however, is the lack of funds.

a. Availability of relevant concepts and measures. The process of identifying performance deficits and setting objectives to correct them requires concepts for analyzing management performance and criteria for measuring it. Program managers and operators perceived that concepts and measures for analyzing and measuring performance are generally unavailable in policing and in the management echelons of public sector agencies. These dual shortcomings of police and management performance measurement systems converge to limit options in analyzing police management performance. This conceptual and measurement problem is compounded, and also caused, by the lack of empirical data on the police manager's function. Although the literature is rich with analyses of the police manager's role and function, nearly all of these are based on surveys. No empirical studies are readily available.5/ The lack of empirical data hinders design of appropriate performance concepts and measures. This makes it difficult to determine rationally what types of management training are germane to particular audiences of police managers.

b. Availability of research techniques and designs. Programs often have restricted research techniques and designs for working with available concepts and measures. Research techniques and designs developed mainly by business, industry, and Federal agencies have not been widely disseminated within criminal justice training programs. This tends to confine the program developer to customary ways of formulating goals, assessing needs, and setting objectives.

c. Availability of adjunct data sources. When the developer forms a strategy for identifying goals and assessing needs, certain adjunct data sources can be useful. These include centralized information about the training histories of the target population, task analyses and management audits from jurisdictions within the target audience, and needs assessments previously performed on the same target population. All these resources can help focus the early steps of program development; however, they are not uniformly available. In many cases, program developers did not even have access to centralized information about the target population's training histories. Where centralized data did exist, they typically showed the number of hours of training received but indicated little about content and outcomes.

5/ Grosman based his conclusions about the police executive's role on informal observation of police executive behavior. Mayo drew upon a more structured and controlled series of observations. An ongoing study, conducted at the University of Southern Illinois at Carbondale with NIJ funds, also involves structured observation. For completed works, see L. A. Mayo, "Analysis of the Role of the Police Chief," (Unpublished doctoral dissertation, American University, 1980), and B. Grosman, Police Command--Decisions and Discretion (Toronto: Macmillan, 1975).

d. Availability of skilled personnel. Training personnel are typically selected more for general administrative abilities and skills in program delivery than for other abilities. Consequently, they often lack technical knowledge needed to conduct the early steps of program development. This technical knowledge can include general knowledge of program development procedures and specific familiarity with the other technical resources described above. Program managers and operators attributed the shortage of skilled in-house personnel and the lack of funds for hiring outside consultants to overall funding and salary constraints.

e. Availability of lead time. For the processes of goal formation, needs assessment, and objective setting to be systematic and deliberate, they must be incorporated into the planning process of the larger organizational environment with reasonable lead time. If programs are scheduled on short notice, these early steps tend to be collapsed or eliminated altogether.

2. Target Population Characteristics. This factor entails the size, stability, and homogeneity of the target audience and the influence these characteristics have on the feasibility and usefulness of early developmental steps.

a. Size of target population. The number of persons in the target population and their geographic dispersion influence the early stages of program development. The more officers in the target population and the wider the geographical area over which they are dispersed, the harder it becomes for the program developer to obtain representative broad inputs to goal formation or to perform a rigorous needs assessment. Programs directed to a national audience are the most vulnerable to this limitation.

b. Stability of target population. The speed and predictability of changes in the target population's dimensions also influence the early developmental stages. Rapid and unpredictable changes in the target population curtail the usefulness of any information collected because any consensus formed will not long remain valid. In this circumstance, allocating program resources to obtain broad inputs systematically and to assess training needs rigorously is like mobilizing an army to capture a ghost.

c. Homogeneity of target population. Differences from one jurisdiction to another in such factors as size, population density, and geographical characteristics shape different management roles and create different training needs. When several jurisdictions have widely discrepant training needs but the resources to operate separate programs for each set of needs are unavailable, program goals tend to reflect both commonalities and differences in need. If objectives recognize and maintain legitimate differences, it can still be difficult to conduct programs targeted toward participant needs. Programs that serve multiple police agencies, varying in size and structure and distributed over a diverse geographic area, consistently noted this problem in program development. One state program noted, "The eastern and western slopes of the state are very different, the one rural and the other urban. We cannot develop a program to satisfy both groups." Some program managers and operators argued that target population homogeneity affects the whole developmental process.

C. FACTORS AFFECTING PROGRAM DESIGN AND IMPLEMENTATION

The five general factors that affect overall program development and two others influence program design and implementation. Program operators most frequently cited legal requirements as limiting their ability to follow a particular developmental path. Direct observation established that pre-existing program materials were of paramount imporance, however. Over and above the five general factors are the two additional factors, instructor availability and trainee availability.

1. Instructor Availability. This factor refers to a program's ability to identify and retain instructional staff suitable for achieving the program's goals and objectives. Instructor availability can be analyzed in terms of the supply of qualified instructors, lead time allowed to obtain instructors, and political and institutional pressures on instructor selection. Based on the national survey, the lack of qualified instructors in a jurisdiction and the lack of lead time to obtain appropriate instructional services have a roughly equivalent, moderate impact on instructor availability. Political and institutional pressures on instructor selection have noticeable but low influence on instructor availability. We suspect that political and institutional pressures actually have greater influence but were underreported, for the same reasons that union influences were underreported. First, they sound like undesirable events to report; second, these influences tend to be subtle, and the typical survey respondent may not be aware of them.

a. Supply of qualified instructors. The absolute number of qualified instructors in an area and the program's ability to identify them are the two key components of instructional staff supply. The number of qualified instructors affects mainly programs outside urban areas, to which physical access is often restricted. In these programs, the use of the immediately available supply of instructors has to be weighed in terms of effectiveness and cost against use of those who might be brought in from outside the jurisdiction. The more critical issue, typically, is identification of the appropriate instructional staff for a program. This can be difficult because the knowledge and skills needed to provide competent instruction depend on the exact model or models by which a program operates. A POST director summed up their difficulty by commenting, "There are an awful lot of people who claim to know something about management." The practice of mixing instructional personnel, often hired on a hit-or-miss basis (described in Chapter Two), reflects problems in identifying qualified instructors and also ambiguity over the model or models by which a program operates.

b. Lead time allowed to obtain instructors. Many instructors work under contract to multiple programs on a part-time basis. A program's capability to project long-term plans affects whether the instructional staff the program wants will have prior commitments. The amount of lead time available to a program to obtain instructors often depends on factors outside its control, including the promotion calendar and funding cycle. The practice of conducting programs contingent on promotions typically means that lead time is extemely short, often no more than a few weeks. Even major national programs, which plan their activities for a year at a time, often have inadequate lead time because the personnel they seek tend to be nationally recognized and make their commitments far in advance.

c. Political and institutional pressures on instructor selection. Though program operators were generally not inclined to discuss these pressures "for the record," several noted them off the record. We later confirmed their remarks by review of program files and instructor interviews. Pressures result in the selection of at least three groups of instructors based on factors unrelated to established qualifications. Friends and acquaintances of those controlling resource allocation are selected to assure funding flow. Local chiefs are selected to ensure a continuing flow of trainees from their departments. Instructors from the host institution are selected to maintain institutional acceptance of the program. In some instances, programs readily accept these "recommendations" for instructor selection because they have been unable to identify appropriate personnel in other ways. These recommendations often seemed to result in effective selection decisions. The presence of strong political or institutional pressures on selection, however, can also close off the program's access to the larger supply of qualified personnel.

2. Trainee Availability. This factor refers to a program's ability to attract trainees with the characteristics and training needs for which the program was designed. Trainee availability can be analyzed in terms of coverage requirements in the agencies served by a program, tuition fund availability, union demands for overtime pay during training, pressures to make greater use of local programs, the adequacy of incentive funds to attract participants, divergence in perceptions of training needs between the program and served agencies, the relative priority placed on training in participating agencies, court challenges to departmental promotions, saturation of the local training audience, misrepresentation of trainee characteristics by the host agency, and system pressures to maximize enrollments to justify a budget. Although we cited a few of these factors earlier, we repeat them here to show their cumulative effect on trainee availability.

Before we describe how these 11 factors influence trainee availability, we can place them in some perspective with results from the national survey. The survey asked about the extent to which certain factors "restricted the availability of trainees for which the program was designed." Based on the national survey, two factors strongly determine trainee availability: coverage requirements in agencies served by the program and the relative priority placed on training in participating agencies. Two factors have moderate influence on trainee availability: saturation of the local training audience and the availability of tuition funds. Four factors exert low influence on trainee availability. These are: pressures to make greater use of local programs, pressures on the program to maximize enrollments to justify its budget, adequacy of incentive funds to attract participants, and union demands for overtime pay during training. The other three factors listed above were not included in the national survey.

a. Police agency coverage requirements. Allowing one or several managers to interrupt the duty schedule to attend training can pose coverage problems in large and small departments alike, especially when fiscal conditions have intensified staffing constraints. In smaller departments, the problem of coverage tends to be greater because there are fewer personnel available to provide coverage. Inability to provide coverage can inhibit departmental willingness to permit managers to take part in training, especially out of state, where

they can less easily be called back. Coverage requirements can also affect whether trainees who attend local programs, in a regional academy or other setting, can take full advantage of training. Some departments, for example, require a manager to serve an eight-hour shift after completing a day-long training session at a local training site. This can affect trainee motivation and in-program performance. Similarly, many trainees remain "on call" during training sessions. This can affect their in-program performance and disrupt program continuity for other trainees. So coverage problems affect not only attendance but also whether trainees in attendance can take full part in training.

b. Tuition fund availability. Tuition funds are especially needed to attend out-of-state programs. Availability of tuition funds for this purpose varies, depending on general departmental funding levels, coverage problems, departmental training priorities, and immediate status of LEAA.

c. Union demands for additional compensation. Union compensation policies sometimes require payment of overtime for part of the time spent in training. These seem to apply mainly when training occurs out of state, on the premise that the training day is longer than eight hours. These policies can deter financially strapped jurisdictions that face coverage problems from giving consideration to training options viewed as marginal or of low priority.

d. Pressures to make greater use of local programs. The system pressure to take residential programs "off campus" or to initiate intrastate regional programs based in local colleges or regional academies tends to draw trainees away from the residential programs. This can result in reduction in the number of sessions offered or in changes in the target population.

e. Adequacy of incentive funds. Pay supplements can provide a powerful impetus for trainee attendance at programs, but a trainee who has obtained the maximum pay incentive in perpetuity will have little extrinsic motivation for attendance. One POST requires annual attendance at training to maintain eligibility for pay supplements. Another POST gives supplements in perpetuity for training and/or advanced education up to a low ceiling amount. This reduces extrinsic motivation for the highly trained or educated officer. The second POST holds the ceiling low, on the apparent assumption that trainees who have "come this far" have generated intrinsic motivation for training attendance.

f. Divergent perceptions of training need. A police agency that sees its training needs as divergent from those of a particular program is likely to send its trainees to other programs more compatible with its own perceptions of needs.

g. Priority placed on training by affected agencies. Training priorities can affect trainee availability in many ways. The most obvious are the frequency and extent to which the department will free up personnel to attend training, provide coverage and tuition support, and make selection decisions on the basis of performance and for purposes of career development. A less obvious effect of training priorities is the extent to which training activities are coordinated with other departmental activities. We saw the effects of inadequate

coordination in a regional academy program, delivered by a major national provider's field division. In this case, the academy's largest target jurisdiction scheduled a promotional exam on short notice in conflict with a training session at the regional academy scheduled a year in advance. Because most trainees were scheduled for the exam, they had a choice: to attend training and defer chances of promotion or not to attend training. Most opted for the promotional exam and delegated responsibility for program attendance to lower ranking personnel, who were not appropriate for the program.

h. Court challenges to departmental promotions. Programs tied to promotions can be affected by court challenges to the promotional process. Several major jurisdictions, as noted earlier, have halted program delivery when EEO suits led to suspension of promotions pending the outcome of litigation.

i. Saturation of the training audience. Several training programs often operate in the same area and compete for the same trainees in a depleting and sometimes nearly exhausted market. Saturation of the local training audience leaves programs with two options. They can cut back offerings or they can accept trainees who do not possess characteristics of the intended audience.

j. Misrepresentation of trainee characteristics by the host agency. To obtain the services of field programs offered by major national providers, host departments or regional academies sometimes "jockey for position," in the words of one program administrator. They do so by deliberately misrepresenting the characteristics of the population to be trained. Because field programs exercise little or no direct control over trainee selection, they have few means to ensure attendance by the population for which the program was designed. Field programs with two characteristics are especially vulnerable to misrepresentation: ones that are renowned and recognized for the quality of their services and ones that are free. Field programs conducted by the FBI National Academy are, consequently, more susceptible to "jockeying" than most field programs.

k. System pressures to maximize enrollments. System pressures to "play a numbers game" to increase budgets and enhance program status often lead to disregard for selection standards. In the absence of performance measures to demonstrate program effectiveness, oversight agencies and program operators tend to measure success by a program's popularity. They view expanded popularity as justification for budget increases. The equation of popularity with success can be seen on the Federal level. In the early seventies, it caused the FBI to focus upon response to an increasing number of police agency requests for assistance at the expense of the Bureau's mandate to stimulate state/local self-sufficiency in training capability. This emphasis on quantity reportedly generated increasing budgets during this period, but it also caused many police agencies to view the FBI as a permanent source of instructional assistance, thereby decreasing their efforts to achieve self-sufficiency.

On the state and local levels, many programs assume that, the larger the audience they can attract, the better will be their bargaining position when re-funding comes into question. One observed program had been locked for years in a battle with the state police agency for control of the state's police training budget. This program was not atypical in placing almost exclusive emphasis

on maximizing the number of enrolled trainees without regard for selection criteria. State and local agencies also often use FBI courses as drawing cards for which they can take credit without using their own resources and sometimes pack them to maximize enrollments.

D. FACTORS AFFECTING THE CONDUCT AND USE OF EVALUATIONS

The five general factors that affect overall program development and three others influence the conduct and use of evaluations. The programs we observed most often cited funding constraints as limiting evaluation capability. It also appeared that the organizational environment often obviated the need for evaluations by reducing their likely usefulness. Over and above the five general factors, these three also affected the conduct and use of evaluation: funding requirements for evaluation, technical resource availability, and resistance to evaluation.

1. Funding Requirements for Evaluation. This factor involves the imposition of an evaluation component upon a grantee by a funding agency as a condition of continued funding. Aside from requirements to conduct examinations and survey the reactions of trainees at the conclusion of training, few programs have been subject to funding requirements of this nature. All the major management training programs currently or previously funded by LEAA have been obliged to meet some type of follow-up evaluation requirement. Required evaluations varied in rigor and typically consisted of follow-up utilization surveys. Although funding requirements have led to more evaluation efforts, they have not shifted the focus of evaluations to performance measurement.

2. Technical Resource Availability. This factor encompasses the availability of personnel, techniques, and other resources needed to conduct evaluations. This factor can be analyzed in terms of availability of relevant concepts and measures, research techniques and designs, skilled personnel, and controls over departmental policy.

a. Availability of relevant concepts and measures. Program managers and operators perceived that concepts for analyzing and measures for assessing performance are inadequate both in policing and in the management echelons of public sector agencies. These dual shortcomings of police and management performance measurement systems converge to limit options in analyzing and measuring police management performance.

b. Availability of research techniques and designs. Programs often have restricted access to research techniques and designs for working with available concepts and measures. Research techniques and designs developed by business, industry, and Federal agencies for evaluating management training programs have not been widely disseminated in criminal justice agencies.

c. Availability of skilled personnel. Staff of most programs do not possess sophisticated evaluation skills and spend most of their time in program delivery and administration. Several observed programs echoed the exaggerated view of one POST director that, "To evaluate a program adequately, the evaluation staff would have to be larger than the training staff. As it is, we are stretched thin."

d. Availability of controls over departmental policy. Rarely does a police agency see training as crucial in a coordinated strategy to upgrade departmental performance. Thus, the program operator has little or no control over changes in departmental policy that can affect variables relevant to an evaluation. One major municipal police agency, for example, has for several years contracted to a management consulting firm recognized worldwide for its leadership training courses. Though the program is managed by the chief's office, the occurrence of changes in management practice unrelated to the program will probably prevent implementation of the program's evaluation component. So, even in the rare instance where a police agency sees management training as essential to an "experiment" in changing departmental practices, like this one, the variables that contribute to change are typically confounded.

3. Resistance to Evaluation. This factor refers to perceptions by any or all parties involved in training that potentially hamper cooperation with evaluation efforts. These views may be valid or invalid in a particular circumstance, but, either way, they place roadblocks before an evaluation effort. This factor can be analyzed in terms of seven propositions:

- The program's expectations about what it intends to accomplish are not clearly defined.
- The program's expectations are implausible because resources are inadequate or ineffectively used.
- There is little or no agreement over what variables ought to be measured.
- Existing data systems are generally unreliable.
- Evaluation costs are out of line with potential uses of evaluations.
- The political burdens or threats posed by evaluations make them infeasible.
- Evaluations are not likely to have much effect on program operations.

Based on the national survey, two perceptions of factors that "limit the amount and quality of program evaluations" predominate. Programs do not conduct more and better evaluations mainly because evaluation costs are out of line with the potential uses of evaluations and because there is little or no agreement over which variables ought to be measured. The other five perceptions that produce resistance to evaluation, often appropriately, are also widely shared. Survey respondents said that all five limit the amount and quality of evaluations to a moderate extent. The factors that have the lowest reported effect on program evaluation are the lack of plausibility in program expectations and the political threats that evaluation can pose.

CONTINUED

1 OF 4

a. Evaluations as based on poorly defined expectations. Some program managers and operators viewed evaluation design as an impossible task because the program's expectations were so ill defined. The operators of the programs we observed generally were not aware of problems of definition, however, and few had seriously considered major changes in their approach to evaluation.

b. Evaluations as based on implausible expectations. Only a few operators of visited programs had any sense that some of their objectives might not be plausible. A reasonably well-defined set of expectations seemed to be a precondition for awareness of plausibility.

c. Evaluations as employing unacceptable measures. Program directors and instructional staff, trainees, oversight and funding agency staff, and evaluators often face difficulties in agreeing upon the relevant measures of program success. This situation often results in a stalemate.

d. Evaluation systems as unreliable. Some program operators seemed to distrust evaluation techniques because the state of the art does not always permit reliable measurement. Some also objected that they could not adequately control evaluation results. One POST director reflected this distrust, saying, "We tried to evaluate our basic training program, but when evaluation results started to show a negative short-term effect on attitudes, we ceased the evaluation." In other words, because he considered the results unacceptable, he blamed the evaluation techniques and criticized evaluations in general for their potential unforeseen results.

e. Evaluations as not cost-effective. Many program operators and managers objected to the classical program evaluations on the basis of cost and utility. These evaluations can be long, drawn-out affairs and can consume large portions of available financial, personnel, and technical resources. The results tend to be of marginal significance to programs because they are not likely to have much effect on program operations. Thus, evaluations are generally seen as an ineffective use of scarce resources.

f. Evaluations as politically threatening. Those supportive of a program often fear evaluations as weapons held by funding agencies and political opponents to "pull the plug" on funding or "blow the program out." In one observed program, a supportive POST council clearly obstructed the evaluation efforts of the SPA based on such fears for program continuity.

g. Evaluations as not usable. For two distinct reasons, program operators and managers viewed evaluations as unusable. First, they often addressed the interests of the evaluator rather than those of the program and were written in unintelligible "consultantese." Second, certain factors inherent to the training situation--such as command structure, training priorities, legal mandates, and the program operator's sense of mission--made it unlikely that persons in authority would be willing to change the program.

* * * *

In this chapter, we have looked at factors external to programs, and largely beyond the program developer's control, that affect how closely program development practices can correspond to any chosen system. The chapter shows how the program developer's options can be ringed by a multitude of forces beyond his control. Often these forces seemingly conspire to make the program developer abrogate his own assumptions. Several POST directors expressed fatalism about the feasibility of systematic program development by chaining together external forces in a vicious circle. One of them went like this: high turnover among police managers prevents utilization of the results of training; low utilization deters the public from approving training expenditures; lack of public approval for training expenditures restricts funds available for management training; lack of funds keeps training on a voluntary basis; the lack of training mandate reduces motivation to attend training and impairs participant attentiveness; inadequate program participation keeps police management capability low; low capability leads to public dissatisfaction with police services; dissatisfaction with police services leads to high turnover, thus completing the circle. Many of the external factors were viewed as reinforcing each other and deterring greater deliberation in program development.

Are we meant to conclude that systematic program development is now beyond the realm of feasibility for most programs? Program managers and operators expressed the consensual view that a high level of correspondence with the industrial model is infeasible now and is unlikely to become more feasible in the near future. With this conclusion, we largely concur. We also stress, however, that most program managers and operators have the capacity now to make isolated changes that will make their programs more manageable. We do not advise making isolated changes just to come closer to the industrial model. Each potential change should be carefully scrutinized to see what it implies for the program's evaluability.

For most program managers and operators, the first step should be to take an inventory of the program as it presently exists--its resources, its activities, and its expectations. Regardless of how systematically the program's parts were assembled, the starting point should be the resources the program now has, the activities in which trainees take part, and the expectations that the program holds for trainees. Once the basic inventory is complete, then it is possible to assess the conditions and practices that hinder effective program management and to identify ways to modify the program usefully. This is the approach that we take in Part 2.

PART 2: EVALUATION AND MANAGEMENT OPTIONS

In the first part of this report, centered around the industrial model of program development, we found ourselves caught in a cross-fire between two armed camps--those who advocate strict compliance with the industrial model and those who sharply criticize the model as unnecessary, infeasible, or both. From our position, we concluded that an altogether different focus would be more productive for most programs. Therefore, we retreated from the hallowed industrial model--still recognizing that there is room for further experimentation with the model's adaptability--and took another tack.

This part of the report explores other ways to think about program evaluation and management by asking four broad questions:

1. Do the varied ingredients that form police management training programs coalesce into a single, unified training model? Or are programs governed by several distinct models of training?

2. What types of obstacles impede both useful program evaluation and effective program management? What can be done to overcome these roadblocks and make programs more evaluable?

3. How can evaluation show whether training makes any difference in later job behavior?

4. What types of measurements could be taken, at least theoretically, to evaluate program success?

Chapter Four explores Question 1 by sorting out the range of elements that go by the name "police management training" into models. Such models could differ in several respects: the inputs or resources needed to make the model run (the types of trainees, the types of instructors, the types of materials that feed into overall goals); the types of activities in which trainees take part; the immediate outcomes (the hoped-for or desired changes in the trainee that are sought within a course); or the expected effect(s) of training on trainee job behavior, the trainee's agency, and even the larger criminal justice system. We identify a total of fourteen distinct models. Some models deal mainly with differences in the substantive material that trainees are expected to learn; others focus more on the implementation process and on the effects that training can have on police agency function. We present each model based on the more detailed descriptions that we made of the programs we visited. We do not endorse certain models as more effective, useful, or valid than others.

Chapter Five looks at the issues of program evaluability in Question 2. The chapter's main purpose is to provide a self-assessment guide, or Evaluability Checklist. We start by listing the seven preliminary questions that all program managers and operators should ask periodically to see how their program could operate more effectively. (Each question represents an evaluability criterion.) We break down each question into subquestions. For each subquestion, we identify conditions or practices ("roadblocks") that make a program less than optimally evaluable. Then, for each roadblock, we specify one or more ways to eliminate it or, at least, make it less problematic. This self-assessment is the chapter's main purpose, but it also assesses the state of program evaluability on a national scale by indicating how frequently programs throughout the country confront each roadblock.

Chapter Six examines Question 3. It describes three evaluation approaches that certain programs may be in a position to try. All three approaches try to see if training makes any difference in later job behavior, each in a different way. The Action Plan Follow-Up Approach determines whether trainees successfully carry out their intentions by having trainees complete a personalized action plan at the end of training and then systematically following up to see what actions they really carry through. The Proficiency Analysis Approach asks whether training improves the trainees' proficiency as managers in ways that their agencies require and can use. This approach uses a pretraining personalized training needs assessment for each trainee, a similar post-training needs assessment (after trainees have returned to their job for a predetermined period), and the comparison of pre- and posttraining needs assessments to see if and how training needs decrease and trainees become more proficient. The Simplified Cost-Benefit Analysis Approach asks whether training is a reasonable investment to make, based on its return to the organization in increased job proficiency. This approach takes the results of the Proficiency Analysis and plugs in cost data (such as salary and benefits) for each trainee to see if any improvements in job proficiency were worth the agency's investment. We have tried out the first two approaches and can say that, from our experience, they often produce useful information.

Appendix 6 addresses Question 4. We outline the key measurement points found in Chapter Four's models and systematically identify potential measures and data collection approaches for each. Most of the variables in the models have not been well measured, often for good reason. Many are extremely difficult to measure, and we have no delusions listing measures will greatly enhance future evaluation efforts. We present these measures and data collection approaches as an appendix rather than as a chapter because they represent a technical reference document.

The three chapters and the technical appendix in Part 2 draw on all the study's data sources. Chapter Four draws primarily on the descriptions that we made of visited programs for the raw material used to construct the models. Also, to a lesser extent, we used expert consultation, the preliminary telephone surveys, and a review of program brochures to construct the models. Chapter Five builds mainly on site-visit observation but also draws on the evaluation literature, preliminary telephone surveys, and the national mail survey. Chapter Six is based on review of the evaluation literature, consultation with experts in

management training evaluation, and our experimentation with two of the three described evaluation approaches. Appendix 6 draws on review of the training evaluation literature and site visit data. Taken collectively, these three chapters and technical appendix have been described by reviewers as a "manual for program management and evaluation."

Four: Program Models

Do police management training programs follow a single model or several different ones? Some program managers and operators seemed to think there is only one model that programs could follow, based on their remarks like, "Management is management," and, "There is only one thing management training can involve--the principles of management." Based on examination of operating programs, however, nothing could be further from the truth. Police management training programs take many forms, varying in the functions they seek to serve and in the means used to achieve chosen ends. They deliberately use widely differing types of resources, do different things with trainees, try to produce different types of changes in trainees during a course, expect and want trainees to try different types of things back on the job, and hope to impact on police agencies and the larger criminal justice system in different ways. Clearly, there is no single model of police management training.

The purpose of this chapter is to lay out the assumptions that underly each police management training model and the intrinsic logic that makes each model work. The models represent actual variants of training. They were derived from program activities that we observed in the field and from discussions with persons responsible for program management and operation. Each is an empirical, inductive model, abstracted from detailed descriptions of observed programs to show the key features they share. No program, however, fully articulated any one model, and each program also mixed several models together.

Do the variations among models reflect mainly the substantive material that trainees are expected to learn or differences in how the implementation and transfer process is conceived? This chapter presents both basic and auxiliary models of police management training. We can distinguish these in two ways. First, the basic models are each tied to particular bodies of substantive management content; the auxiliary models express no content and may be appended to nearly any basic model. Second, the basic models focus on the processes for transmitting a body of knowledge along with related skills and attitudes to trainees; the auxiliary models focus on the department- or system-level impacts and the processes for generating them. The first three parts of this chapter describe basic models: three compliance, three prescriptive, and two participative models. The fourth section presents six auxiliary models. The fifth closes the chapter by turning attention to how programs mix models and to the effects of this interplay.

The description for each model contains four elements: overview, underlying assumptions, flowchart, and narrative accompaniment to the flowchart. The overview briefly highlights the distinguishing features of a model. The underlying assumptions express beliefs about the dimensions of the problem giving rise to the need for training and about the feasible solutions to these problems. The flowchart and its narrative accompaniment show how a model works.

Each flowchart has five columns, showing how inputs, processes, in-program outcomes, short-term impacts, and long-term impacts are presumed to be causally linked.

- Inputs are those resources a model specifies as needed for its processes to work.
- Processes include the formal and informal activities a model specifies for achieving outcomes and impacts.
- In-program outcomes express changes in learning and behavior that a trainee is expected to display within the program before returning to the work environment.
- Short-term or first-level impacts show how in-program outcomes are expected to transfer to the work setting. (They deal mainly with effects on individual performance.)
- Long-term or second-level impacts show how the short-term impacts are expected to produce cumulative results on a larger, departmental- or system-level scale.

Each flowchart captures the critical variables for a given model. For readability, they do not include all the causal lines that would be needed to pinpoint the links within and between columns in order to express how variables within a column build on each other, as well as the full range of reciprocal, interactive, and feedback relationships. The flowchart and its narrative accompaniment express the testable assumptions for a given model. The narrative accompaniment for each flowchart simply restates the logic of the flowchart.

The models can be useful both in program management and evaluation. They can be used to identify the models that a given program follows; to identify the activities that a program ought to conduct, to be consistent with its own objectives; to clarify how a course's substantive influence is meant to impact upon the trainee's work environment; to identify ways that a program may need to clarify it expectations, to be more internally consistent; and to identify variables appropriate for evaluation. In this way, the models can help to sort out the complex management issues that arise in coordinating staff, in stating objectives clearly around central themes, and in mobilizing available resources toward program goals. They can also help to isolate the outcomes and impacts that a program seeks to foster and that would, consequently, be the best measures of the program's success.

Exhibit 14 contains an orientation chart designed to provide a quick overview of the fourteen models. This chart should make it easy to find the models of interest for a particular program.

A. COMPLIANCE MODELS

The compliance models look upon training primarily as a mechanism for establishing and maintaining control and "good communications" among managers in a department. This section contains three compliance models: pre-service and

EXHIBIT 14

NEP/Police Management Training

ORIENTATION CHART TO 14 MODELS

Model	Exhibit Number	Training Activities	Workplace Results
COMPLIANCE MODELS			
Pre-Service and Initiatory	15	Orients newly promoted managers to their new duties and responsibilities.	Trainees perform their duties in compliance with regulations and their units are cohesive.
Refresher and Update	16	Reviews minimum required managerial duties and updates on changes.	Trainees continue to comply with regulations and their units maintain cohesion.
Corrective	17	Reasserts departmental control to correct specified deficiences.	Trainees renew their compliance with regulations and their units become more cohesive.
PRESCRIPTIVE MODELS			
Systematized Policing	18	Challenges trial-and-error decision-making and applies principles derived from business to law enforcement.	Trainees accept and use business principles.
State-of-the-Art	19	Circulates information about innovative practices that police managers have proved effective.	Trainees are less likely to "reinvent the wheel" and more inclined to experiment with tested police innovations.
Adaptation	20	Outlines adjustments needed to comply with changed Federal, state, and local regulations.	Trainees identify practices requiring adaptation and implement new practices that comply with regulations.
PARTICIPATIVE MODELS			
Non-Experiential Participative	21	Explains how to meet the needs of agency personnel for self-actualization and a decision-making role.	Trainees apply participative principles in limited areas.
Experiential Participative	22	Combines explanations of participative concepts with structured experiences in how participative management systems operate.	Trainees assimilate participative experiences, gradually change their attitudes, and apply participative principles.
AUXILIARY MODELS			
"Greasing the Skids"	23	Provides amenities inside and outside the classroom to ensure that trainees go home rested, uplifted, and satisfied.	Trainees obtain boosts in morale and performance and encourage other officers to disply exemplary behavior.
Certification	24	Enforces minimum knowledge standards and provides incentives for course completion.	Substandard managers are weeded out, successful trainees show more interest in career development, more qualified individuals are attracted to law enforcement, and turnover is reduced.
Network	25	Structures intense and extended interaction among trainees during classroom and off-time hours in a residential setting.	Lasting personal acquaintances among trainees perpetuate course effects through assistance to "alumni" in problem-solving and career development and through general reinforcement of course contents.
Anointing	26	Recognizes officers tagged for big promotions by sponsoring attendance at nationally recognized training programs.	New trainee credentials and acquaintances enhance stature and career progression; status of "alumni" reinforces the program's reputation.
Departmental Decision-Making	27	Uses training as a two-way communications vehicle, providing line managers with information about a pending decision and with an opportunity to shape the decision through critical feedback.	Agency makes an informed decision about whether and how to proceed; personnel are less resistant to change and implementation of the decision is smoother.
Critical Mass	28	Promotes large-scale organizational change and tries to create the preconditions for it.	Change-oriented graduates advertise the program, assist other graduates, and lay the foundation for large-scale change. Eventually, there are enough similarly attuned managers to carry out change and enough senior officials to initiate and oversee it.

initiatory, refresher and update, and corrective models. The three compliance models are occasionally combined but are seldom found with prescriptive models because compliance models emphasize the restrictions on managerial discretion formed by the legal/procedural/administrative net in which managers must operate, whereas prescriptive models call for the fuller recognition of managerial options and the exercise of managerial discretion. Because of their control orientation, compliance programs are usually conducted in single jurisdictions solely for trainees from that jurisdiction.

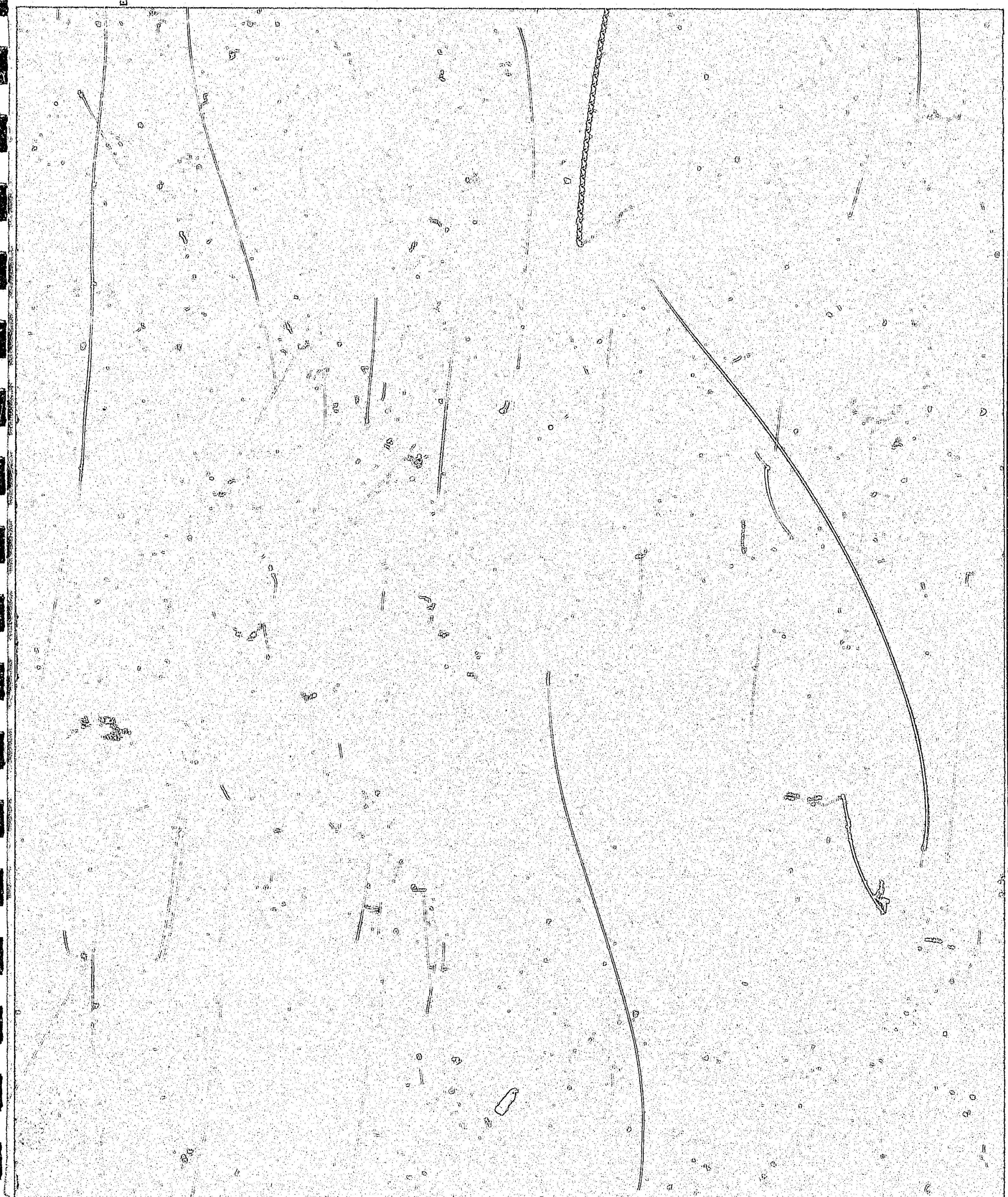
1. Pre-Service and Initiatory Model. The pre-service compliance model orients newly promoted trainees to required duties, contexts within which those duties must be performed, and approaches recommended for the conduct of duties. This orientation is expected to allow trainees to gain confidence and a fresh start in their new positions and avoid forming bad work habits. The pre-service model is based on the following underlying assumptions, three of which it shares with the two other compliance models. These assumptions were outlined by program instructors and senior department officials:

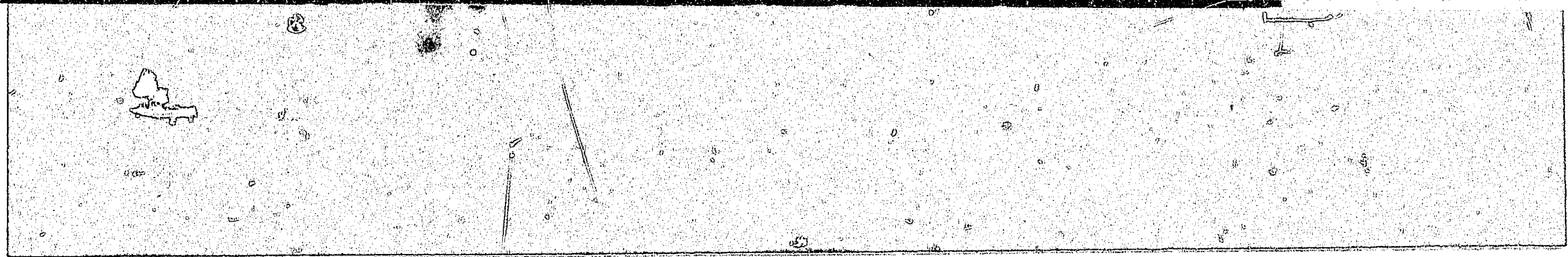
- Executive control over management and line personnel is needed to accomplish departmental objectives and coordinate departmental activities. The implementation of standard policies and procedures is indispensable to control maintenance.
- The need for executive control, the complexity of police organizations, and recent law enforcement legislation have forced police managers to operate within ever-tightening legal, administrative, and procedural limitations.
- A manager's range of potential duties varies widely from rank to rank.
- Progression in management rank and responsibility involves increasing contacts with other departmental divisions and with other agencies that interface with the department.
- Departmental control can be furthered by standardized orientation of new management promotees to their new duties and responsibilities.
- Pre-service training can rely on existing control mechanisms to provide trainee incentives to accept and implement training contents.

Exhibit 15 shows how the pre-service compliance model works.

a. Inputs. The model specifies four inputs: recently promoted managers, senior departmental officials as instructors, departmental training staff, and departmental training facilities.

b. Processes. Six major processes occur during the course of program activities. Trainees are presented with an overview of departmental policies, operations, and interfaces with other criminal justice agencies. They also receive a description and analysis of their minimum required duties and responsibilities. The broad, non-rank-specific legal, administrative, and procedural scope of trainee duties is delineated. Trainees receive a description





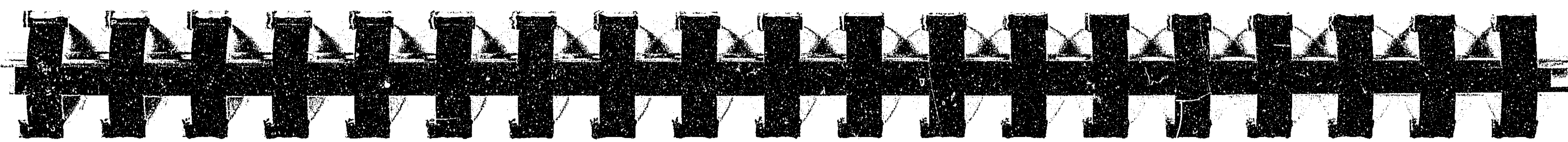
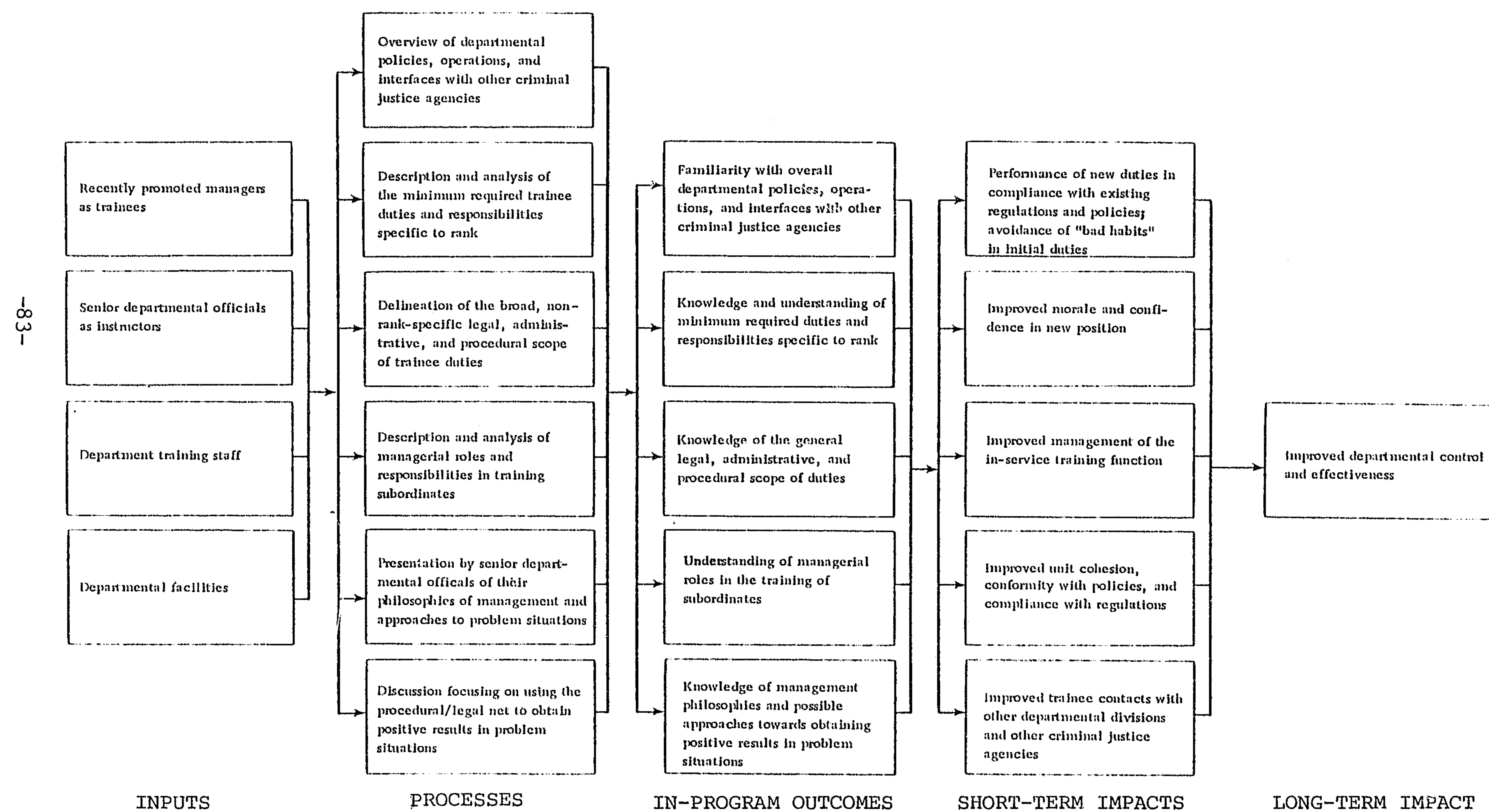


EXHIBIT 15

NEP/Police Management Training

PRE-SERVICE AND INITIATORY COMPLIANCE MODEL



and analysis of managerial roles and responsibilities in training of subordinates. Throughout the program, departmental officials present their philosophies of management and their approaches to problem situations. Finally, instructors and trainees discuss how to use the legal and procedural framework to obtain positive results.

c. In-program outcomes. Five in-program outcomes should result. Trainees become familiar with overall departmental policies, operations, and interfaces with other criminal justice agencies. They know and understand their minimum required duties and responsibilities, as well as the general legal, administrative, and procedural scope of their duties. They also understand managerial roles and responsibilities in the training of subordinates. They are familiar with departmental management philosophies and with approaches to obtain positive results.

d. Impacts. This model should produce five short-term impacts. Trainees perform their new duties in compliance with departmental policies and regulations, thereby avoiding bad work habits. Trainee morale and confidence improve, as does management of the in-service training function. On the unit level, there is greater cohesion, conformity with policies, and compliance with regulations. In addition, trainees strengthen their contacts with other departmental divisions and with other criminal justice agencies. The ultimate impact is improved departmental control and effectiveness.

2. Refresher and Update Model. The refresher and update compliance model recognizes that experienced managers can inadvertently stray from the proper performance of their duties and that the scope and limitations of those duties can change considerably over time. Consequently, to maintain continued compliance with policies and procedures and to ensure departmental control, it sees periodic review of minimum required managerial duties and update on changes as essential. The refresher and update model is occasionally mixed with the pre-service model in departments with multi-level command structures that share common types of management responsibilities. It builds on the following underlying assumptions as stated by program instructors and senior departmental officials:

- Executive control over management and line personnel is needed to accomplish departmental objectives and coordinate departmental activities. Department-wide implementation of policies and procedures is indispensable to maintain this control.

- The need for executive control, the complexity of police organizations, and recent law enforcement legislation have forced police managers to operate within ever-tightening legal, administrative, and procedural limitations.

- Due to rapid organizational change and the variety of situations faced by each police manager, experienced managers need periodic in-service training to refresh prior training and obtain updated information on current duties and responsibilities.

- Control can be furthered by standardized refresher training within rank levels.

- In-service training can rely on existing control mechanisms to provide trainee incentives to accept and implement what they learn.

Exhibit 16 shows how the refresher and update compliance model works.

a. Inputs. There are four inputs to the model: experienced managers of a similar rank or position, senior departmental officials as instructors, departmental training staff, and departmental training facilities.

b. Processes. Eight major in-program processes occur. Trainees review current departmental policies, operations, and interfaces with other criminal justice agencies. Major changes in department policies, operations, and interfaces are described and analyzed. Trainees receive both a review of their minimum required duties and responsibilities and a description and analysis of major changes in those duties and responsibilities. They are provided a review of the broad, non-rank-specific legal, administrative, and procedural scope of duties, along with a description and analysis of major changes in their scope. Senior departmental officials reaffirm their previously stated management philosophies or adapt them to fit major changes affecting trainees. Trainees and instructors also discuss experiences and problems in the conduct of duties.

c. In-program outcomes. Five in-program outcomes may be anticipated. As they leave the program, trainees understand current departmental policies, operations, and interfaces with other criminal justice agencies. They know and understand their minimum required duties and responsibilities and the legal, administrative, and procedural scope of their duties. They better understand departmental philosophies of management and departmental approaches to problem situations. In a broader sense, they know more about management problems and their solutions.

d. Impacts. Two major short-term impacts result from the program. Trainees exhibit continued improved performance of duties in compliance with policies and regulations. Their units show improved cohesion, conformity with departmental policies, and compliance with regulations. The ultimate impact is improved departmental control and effectiveness.

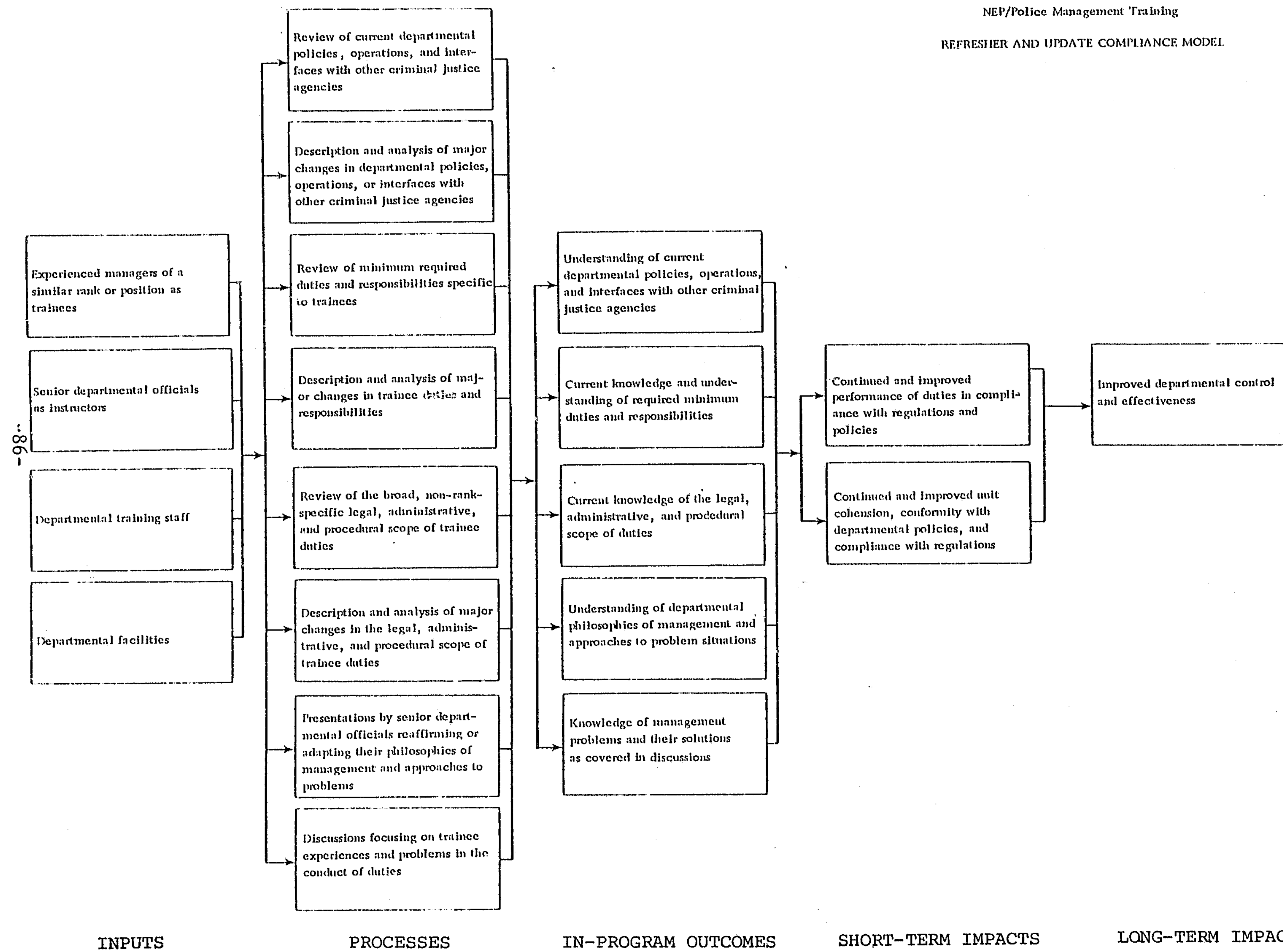
3. Corrective Model. The corrective model reasserts departmental control over managers and corrects specific identified or perceived performance deficiencies common to a group of trainees. It is occasionally combined with the refresher and update model and could even be combined with the pre-service model. It is based on the following underlying assumptions, as articulated by program instructors and senior departmental officials:

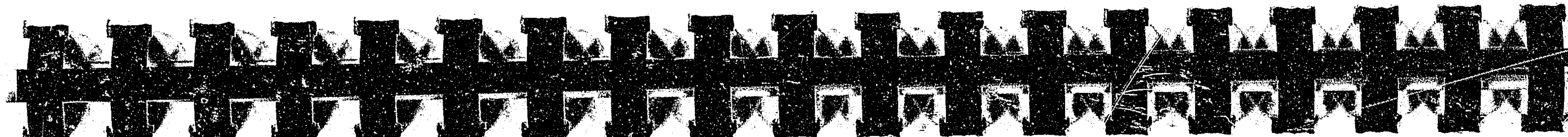
- Executive control over management and line personnel is needed to accomplish departmental objectives and coordinate departmental activities. The implementation of standard policies and procedures is indispensible to maintain control.

EXHIBIT 16

NEP/Police Management Training

REFRESHER AND UPDATE COMPLIANCE MODEL





- The need for executive control, the complexity of police organizations and recent law enforcement legislation have forced police managers to operate within ever-tightening legal, administrative, and procedural limitations.

- Managers can develop performance deficiencies from inadequate or misunderstood information about their duties or from bad habits developed during efforts to implement policy.

- In-service training, supplemented by existing departmental controls, can correct identified performance deficiencies.

Exhibit 17 shows how the corrective compliance model works.

a. Inputs. There are four inputs to the model: experienced managers with demonstrated or presumed common performance deficiencies, senior departmental officials as instructors, departmental training staff, and departmental training facilities.

b. Processes. Five major processes occur during a program. Instructors delineate the specific and general performance deficiencies to be covered. Relevant performance standards, duties, and responsibilities are described and analyzed. The reasons for deficiencies are analyzed, and trainees receive information on how to correct their deficiencies. Instructors and trainees discuss problems in their remediation.

c. In-program outcomes. Three in-program outcomes result from these processes. Trainees understand the reasons and remedies for their performance deficiencies. They better understand their required duties and responsibilities. They are motivated to correct their deficiencies.

d. Impacts. Two short-term impacts may be expected: renewed individual compliance with regulations and departmental policies and improved unit cohesion, conformity with department policies, and compliance with regulations. The ultimate impact of the program, as in the other compliance models, is improved departmental control and effectiveness.

B. PRESCRIPTIVE MODELS

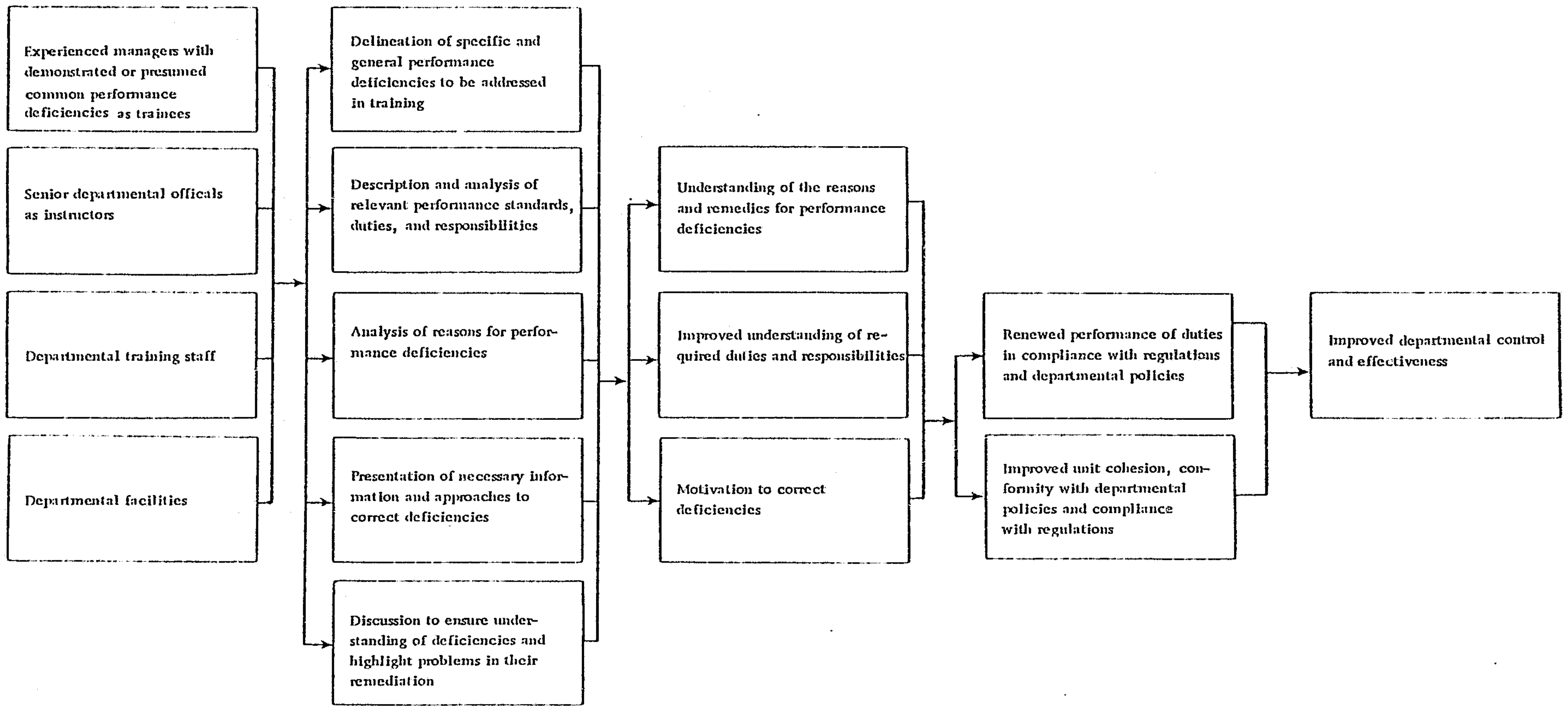
The prescriptive models communicate a body of accumulated knowledge that has broad implications for police management practice and that draws upon experiences of the business community, the experiences of other police managers, and the rulings of regulatory agencies. This section contains three prescriptive models--systematized policing, state-of-the-art, and adaptation models--which are often intermixed, are occasionally found with participative models, and can be linked to any auxiliary model.

1. Systematized Policing Model. The systematized policing model is designed to improve the quality of law enforcement management through the application of principles and practices that have been proven effective in business and industry. The model sees these principles and practices as a stimulus for

EXHIBIT 17

NEP/Police Management Training

CORRECTIVE COMPLIANCE MODEL

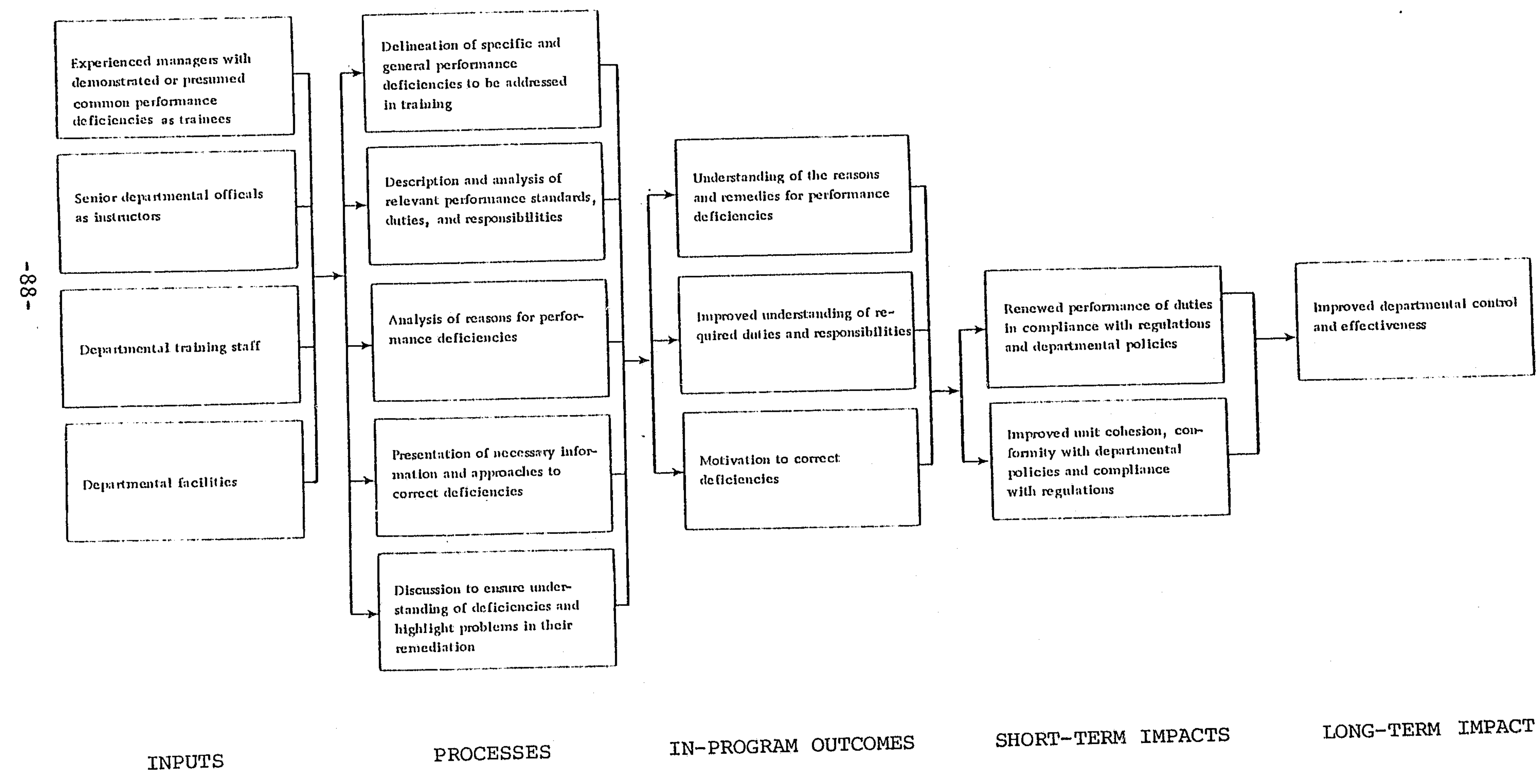


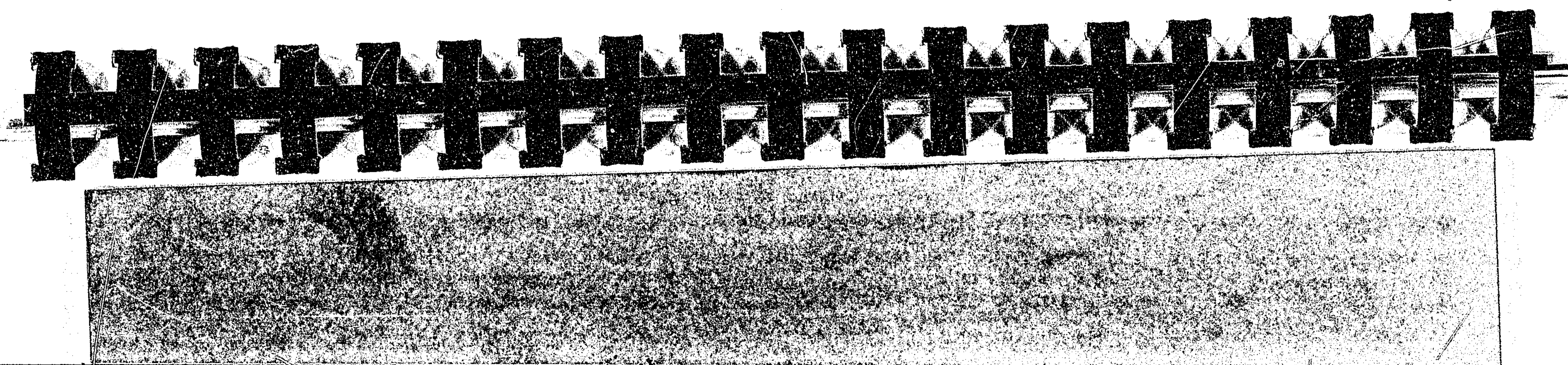
INPUTS | PROCESSES | IN-PROGRAM OUTCOMES | SHORT-TERM IMPACTS | LONG-TERM IMPACT

EXHIBIT 17

NEP/Police Management Training

CORRECTIVE COMPLIANCE MODEL





rational managerial decision-making and, thus, as a challenge to trial-and-error methods common to traditional police management. It treats police agencies like business and industrial organizations, capable of management by similar methods. The claim of similarity is not uniformly accepted, so this model is often mixed with the state-of-the-art or the adaptation models. The systematized policing model is based on the following underlying assumptions outlined by program instructors, administrators, and oversight agency personnel:

- Police agencies' managers have been less than optimally effective because they lacked knowledge of and skills in modern management principles and practices. They have managed their agencies through trial and error, often running an idea up the flagpole to see if it works.

- A time-tested and proven body of management theory and practice, thoroughly implemented by private industry and the military and applicable to private- and public-sector agencies alike, has been available for some time.

- Police managers have generally fishladdered their way through the ranks without instruction in management, and they feel threatened by the introduction of unfamiliar principles and practices.

- Modern management practices, already proved effective at systematizing policing, warrant the consideration of police managers.

Exhibit 18 shows how the systematized policing model works.

a. Inputs. The model specifies only two inputs: managers unfamiliar with modern management theory and practice and instructors well versed in both of these.

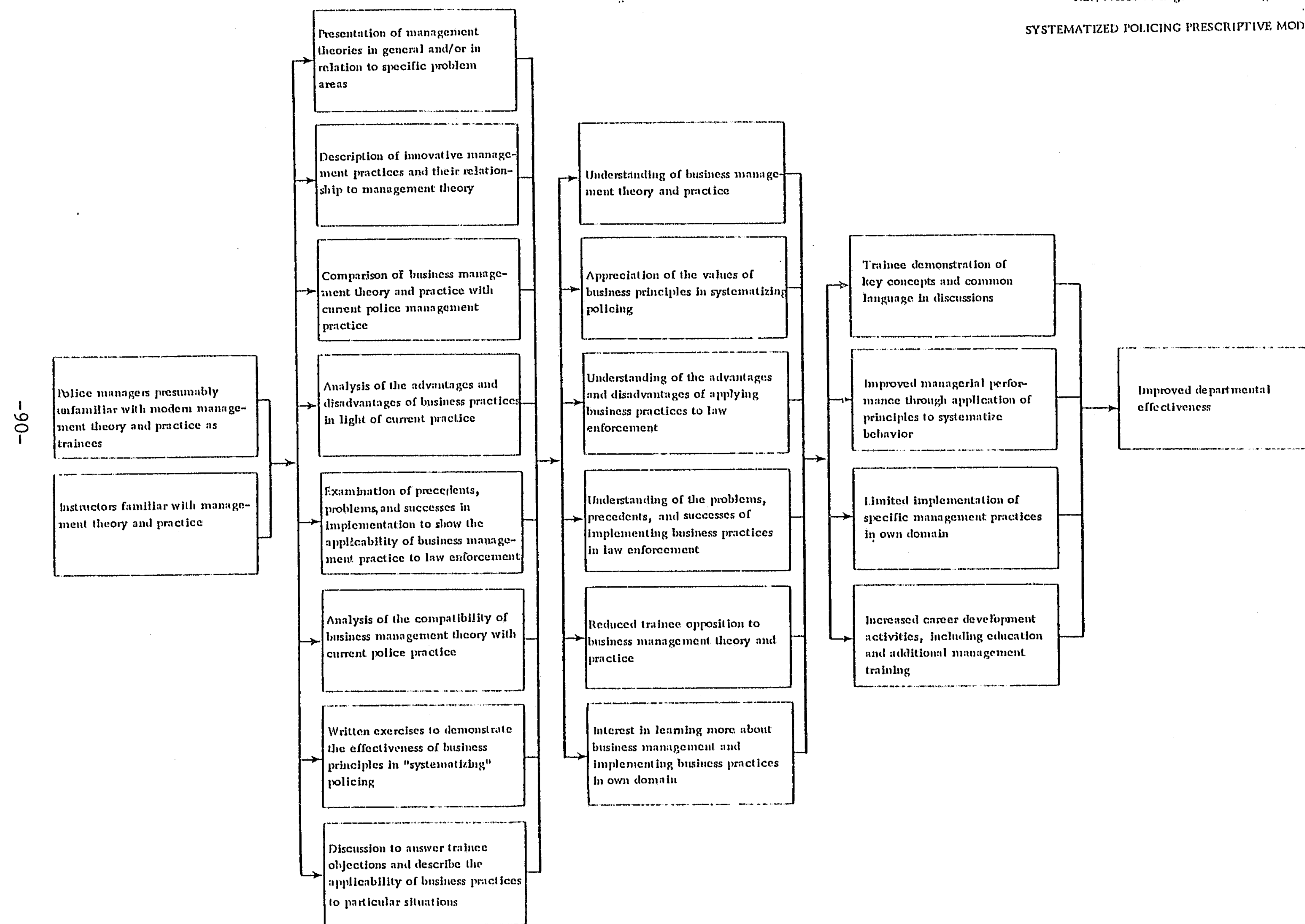
b. Processes. Seven major processes occur during the program. Trainees are presented with management theories, in general and in specific problem areas. This presentation is followed by a description of innovative management practices and their relationship to theory. This theory and practice are compared with current police management practice. The advantages and disadvantages are weighed. Trainees and instructors examine the precedents, problems, and successes of past attempts to implement innovative management practices in law enforcement. Trainees also take written exercises that illustrate the effectiveness of management principles in systematizing policing. Class discussions answer trainee questions and objections and describe the application of management theory and practice to specific areas of interest to trainees.

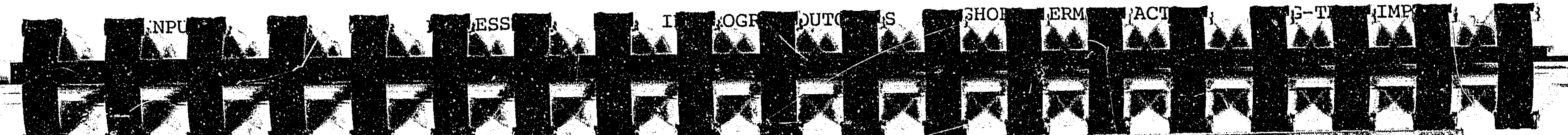
c. In-program outcomes. Six in-program outcomes occur. Trainees generally understand business management theory and practice. They appreciate the value of management principles in systematizing policing. They understand the advantages and disadvantages of applying innovative management practices to law enforcement. They also understand the problems, precedents, and successes of past implementation attempts. Trainee opposition and resistance to business management theory and practice are reduced. Finally, trainees show their interest in learning more about classroom materials and in transferring new ideas to their own departments.

EXHIBIT 18

NEP/Police Management Training

SYSTEMATIZED POLICING PRESCRIPTIVE MODEL





d. Impacts. As a result of these outcomes, four short-term impacts can be anticipated. Trainees demonstrate their knowledge of key concepts and the common language of management in post-program discussions with other managers. Their individual performance improves through the application of program principles in systematized job behavior. Limited implementation of specific management practices occurs within trainee departments. Trainees engage in increased career development activities, including education and additional management training. The ultimate result of these short-term impacts is improved departmental effectiveness.

2. State-of-the-art Model. The state-of-the-art model tries to improve the quality of police management through the circulation of information on the most current, innovative, and effective practices in use by police managers. In this model, police management is a unique profession that cannot be effectively organized on principles borrowed from business or industry; police managers face problems peculiar to police agencies and, thus, benefit most from knowing what actions other police managers have found most effective. This knowledge helps avoid "reinventing the wheel." Managers so informed will also not be inclined to disregard practices that have proved effective in other departments. The state-of-the-art model is based on the following underlying assumptions specified by program instructors and administrators and by oversight and funding agencies:

- Public-sector agencies, including police agencies, differ greatly from private-sector organizations in environment, goals, objectives, and operations, limiting applicability of principles and practices developed by business and industry.

- Because police operations encompass varied duties and diverse restrictions, no single body of knowledge constitutes optimal police management practice.

- Police managers have traditionally operated by trial and error because of inadequate information about effective strategies. This insularity and lack of information-sharing have resulted in departmental ineffectiveness, inefficiency, and duplication of efforts of managers in other agencies.

- Effective police managers should be able to identify and implement current innovative and successful practices from other police agencies, including the most current results of major police management research and development projects.

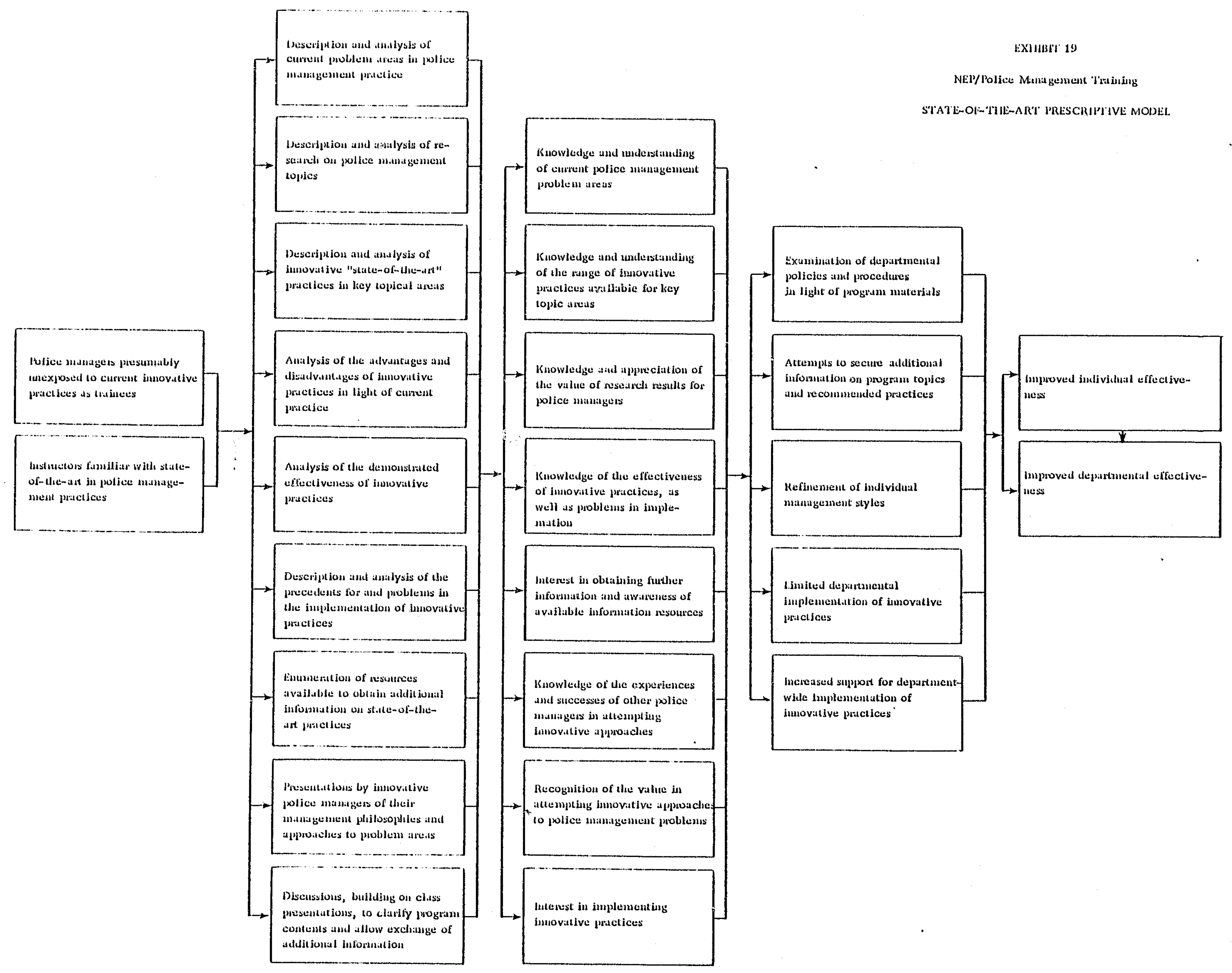
Exhibit 19 shows how the state-of-the-art model works.

a. Inputs. This model specifies only two inputs: police managers who have not been exposed to current innovative practices in police management and instructors who are well versed in state-of-the-art police management practices.

EXHIBIT 19

NEP/Police Management Training

STATE-OF-THE-ART PRESCRIPTIVE MODEL



b. Processes. Nine major in-program processes occur. Current problem areas in police management are described and analyzed. The current state of research on police management topics is described. Current and innovative state-of-the-art practices in police management are presented to trainees and analyzed. The advantages and disadvantages of these practices are weighed in comparison with current practice. Evidence of the demonstrated effectiveness of these innovative practices is presented and analyzed. The problems and precedents for the implementation of these practices are described. The resources available to trainees for obtaining additional information on innovative practices are enumerated. Innovative police managers present their management philosophies and approaches to problem areas. Finally, discussion clarifies program contents and allows the exchange of additional information.

c. In-program outcomes. Eight in-program outcomes are expected. Trainees show improved knowledge of current police management problem areas. They have greater knowledge of the range of innovative responses to those problems. Trainees better appreciate the value of research in police management. They also know more about the effectiveness of innovative practices and precedents and problems in their implementation. Trainees show interest in obtaining further information on innov.tive practices and are aware of the resources available for obtaining it. They know more about the experiences, problems, and successes of other managers in implementing innovative practices; recognize the value of attempts at innovation; and are more interested in replicating innovative practices within their departments.

d. Impacts. The in-program outcomes lead to five short-term impacts. Trainees start to examine departmental policies, procedures, and operations in light of innovative practices described in the program. They try to secure additional information on program topics and recommended practices. The program also results in refinement of individual management styles, limited implementation of innovative practices in trainees' departments, and increased support for department-wide implementation of innovative practices. The ultimate impacts of the program are improved individual and departmental effectiveness.

3. Adaptation Model. The adaptation model focuses on the need for police agency compliance with changes in Federal and state laws, regulations, and policies. This can require overhaul of entire divisions or even comprehensive revision of personnel structure, including recruit selection, promotions, and training, even where traditional organizational structures have seemed effective. The adaptation model outlines the management adjustments necessary to respond to changing legal, procedural, and policy constraints. It is often combined with the two other prescriptive models because the practices it recommends to secure compliance have proved successful in improving agency effectiveness. The adaptation model is based on the following assumptions spelled out by instructors, program administrators, and oversight agencies.

- Changes in Federal, state, and local regulations have obliged police managers to change management practices to remain in compliance with regulations.

- The recent austerity in police agency budgets has forced police managers to allocate existing resources more efficiently to maintain current performance levels.

- To operate effectively, police managers must know about the constraints of changing laws, regulations, and policies, as well as the optimal alternatives for responding to them.

Exhibit 20 shows how the adaption model works.

a. Inputs. The model specifies two inputs: police managers unfamiliar with changed state and Federal regulations and instructors well-versed in legal changes and in effective ways to respond to them.

b. Processes. Six major processes occur during a program. Changes in state and Federal laws, regulations, and procedures affecting management are described and analyzed. Current management practices that are no longer acceptable in light of these changes are identified. Trainees are presented with a description and analysis of the adaptations that are necessary to come into compliance with changed laws, regulations, and policies. The practices that can bring about successful adaptation are described and analyzed. The advantages and disadvantages of recommended practices are analyzed in comparison with current practice. The problems that other managers have encountered in implementing similar practices are presented. Finally, trainees participate in class discussions that justify the need for change, clarify management options for adapting to change, and allow the exchange of further information.

c. In-program outcomes. These processes lead to four in-program outcomes. Trainees understand the changes in laws, regulations, and policies that necessitate adaptation. They also understand the need for change and can identify current management practices no longer acceptable. They are familiar with practices that permit successful adaptation to change. Finally, they know about the problems that other managers encountered in implementing similiar adaptions.

d. Impacts. These outcomes lead to two short-term impacts. Trainees have an increased ability to recognize and analyze situations requiring adaptation. They also make efforts to implement recommended adaptations. The program's ultimate impacts are successful departmental adaptation and maintained or improved departmental effectiveness.

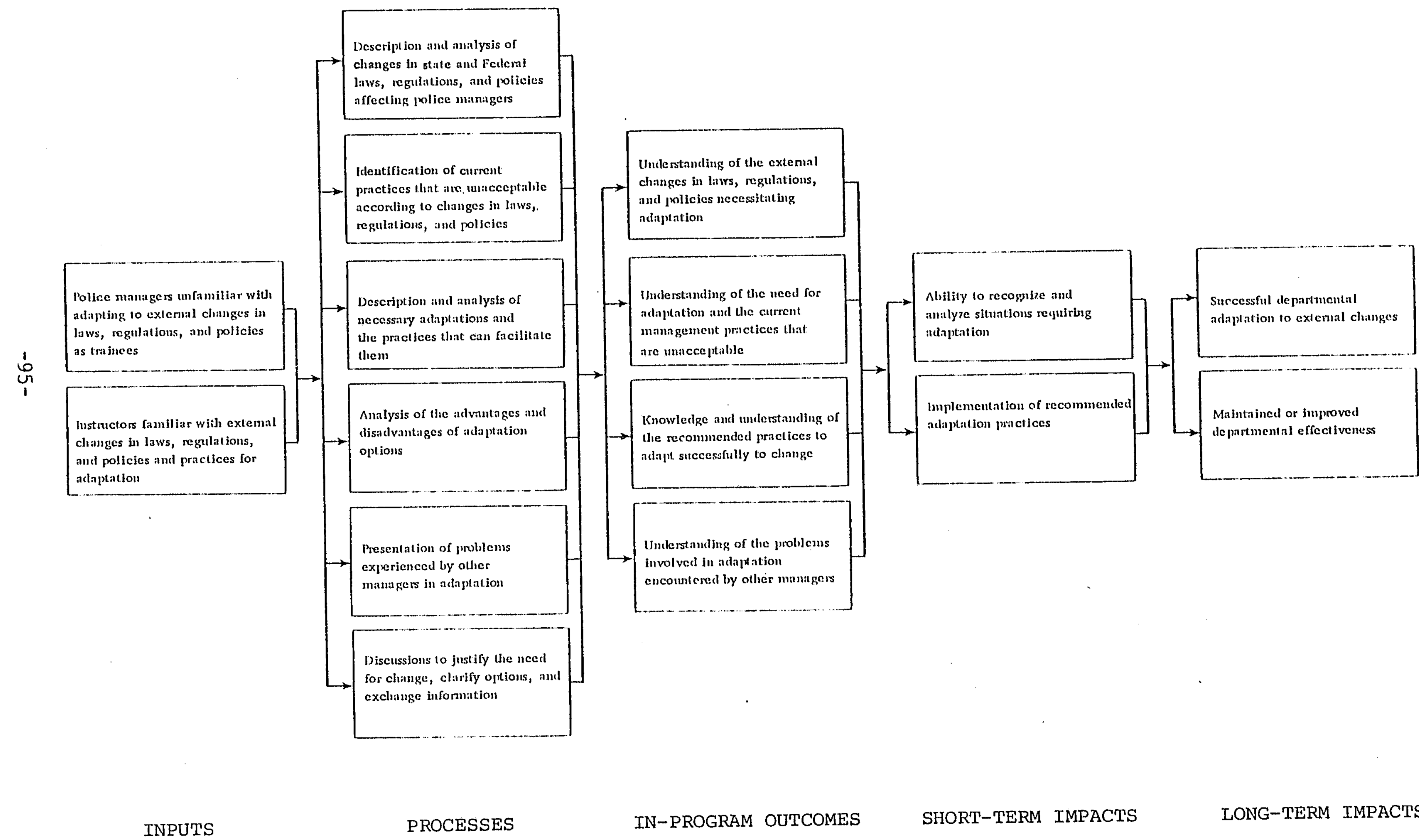
C. PARTICIPATIVE MODELS

Training in a participative model shows police managers how to meet the needs of departmental personnel for self-actualization and for a role in decision-making. This section contains two participative models: non-experiential and experiential. Both stress the need to train managers to understand participative management principles and their operationalization in MBO. They sharply challenge the authoritarian assumptions of traditional police management systems and scientific management systems' lack of appreciation for participative needs. Technically, these are submodels of the systematized policing model derived from business management theory and practice. We consider the participative models separately because of their particular emphases on changed managerial attitudes and styles and on the need to alter organizational structures. The non-experiential model is really an intrinsic part of the experiential model as expressed here. The participative models are occasionally mixed with prescriptive models and can be linked to any auxiliary model.

EXHIBIT 20

NEP/Police Management Training

ADAPTATION PRESCRIPTIVE MODEL



1. Non-experiential Participative Model. The non-experiential participative model uses traditional instructional methods such as lectures, discussions, and case study exercises. It is based on the following underlying assumptions specified by administrators and instructors in participative programs:

- Police managers have been less than optimally effective because they lacked modern management principles and practices and managed largely by trial and error.
- Because traditional and scientific police management systems have overlooked needs for self-actualization and participative decision-making, police personnel have not been consistently well motivated or committed to organizational goals.
- Scientific management principles developed by industry do not allow sufficient recognition of the human element.
- Participative management, usually operationalized as MBO, has significant advantages over traditional and scientific police management systems: it helps generate commitment and motivation, improves departmental teamwork and communication flows, and improves departmental effectiveness.
- MBO is effective for productivity measurement because it requires managers to structure their activities in terms of clearly defined and measurable objectives.
- Because participative management principles challenge traditional authoritarian police management approaches, it is essential to reexamine managerial values and attitudes first.

Exhibit 21 shows how the non-experiential participative model works.

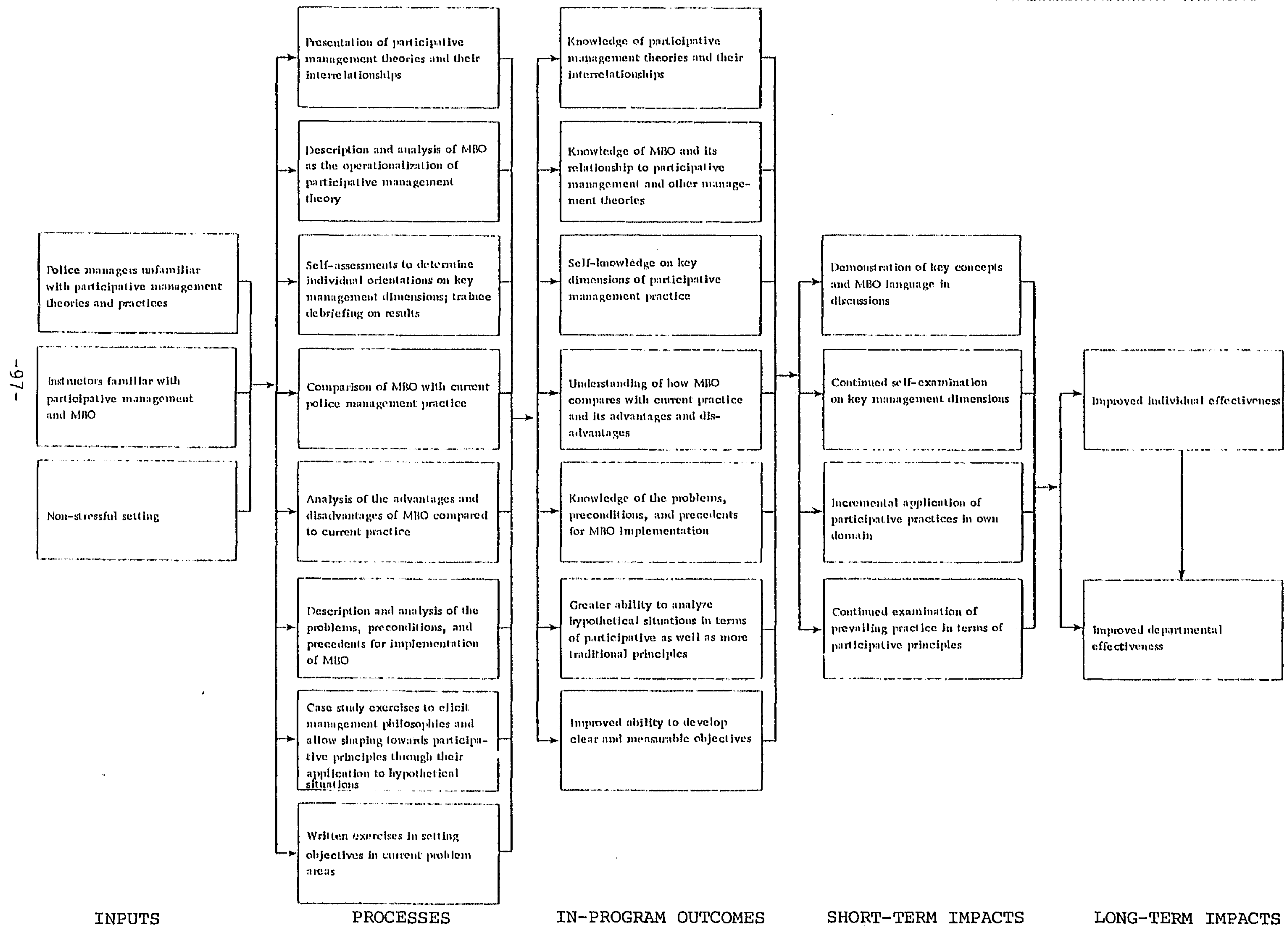
a. Inputs. There are three inputs to the model: managers unfamiliar with participative management theories and practices, instructors well versed in these theories and practices, and a non-stressful setting.

b. Processes. Eight major processes occur during a program. Participative management theories and their interrelationships are presented. MBO is described and analyzed as an operationalized form of participative management. Trainees take self-assessment exercises to determine their individual attitudes and orientations on key management dimensions, followed by a debriefing on assessment results. MBO is compared with current police management practice, and the advantages and disadvantages of MBO in relation to it are analyzed. Trainees are presented with a description and analysis of the problems, preconditions, and precedents for the implementation of MBO. Finally, trainees take part in written case study exercises that elicit trainee management philosophy and allow its shaping toward participative principles through application to hypothetical situations. They also take written exercises in setting objectives for personal problem areas.

EXHIBIT 21

NEP/Police Management Training

NON-EXPERIENTIAL PARTICIPATIVE MODEL



c. In-program outcomes. The processes above lead to seven in-program outcomes. Trainees leave the program with knowledge of participative management theories and their interrelationships. They are familiar with MBO and its relationship to participative and other management theories. They also know more about themselves on key dimensions of participative management practice. They understand how MBO compares with current management practice and also the problems, preconditions, and precedents for MBO's successful implementation. They are better able to analyze management situations in terms of participative and other management principles and to develop clear and measurable objectives.

d. Impacts. The in-program outcomes above produce four short-term impacts. Trainees are able to use participative management concepts and MBO vocabulary in discussions on their jobs. They continue to examine their individual management attitudes and orientations. Along with this self-examination, they incrementally apply participative practices and continue to examine management situations in terms of participative and other management principles. The ultimate impacts of the program are improved individual and departmental effectiveness.

2. Experiential Participative Model. The experiential participative model combines traditional instructional techniques with active individual exercises, structured competitive group experiences, and simulation exercises. Experiential programs use these techniques to stimulate trainee assimilation and internalization of experiences with alternative management systems. This is expected to increase the chances of the attitude change essential to participative management. The experiential participative model is based on seven underlying assumptions, the last of which distinguishes it from the non-experiential model. The following were outlined by instructors and administrators in experiential programs:

- Police managers have been less than optimally effective because they lacked modern management principles and practices and managed largely through trial and error.

- Because traditional and scientific police management systems have overlooked needs for self-actualization and participative decision-making, police personnel have not been consistently well motivated or committed to organizational goals.

- Scientific management principles developed by industry do not allow sufficient recognition of the human element.

- Participative management, usually operationalized as MBO, has significant advantages over traditional and scientific police management systems: it helps generate commitment and motivation, improves departmental teamwork and communication flows, and improves departmental effectiveness.

- MBO is effective for productivity measurement because it requires managers to structure their activities in terms of clearly defined and measurable objectives.

- Because participative management principles challenge traditional authoritarian police management approaches, it is essential to reexamine managerial values and attitudes first.

- So that managers can see themselves and their roles from the perspective needed for participative management, it is important to structure learning experiences that demonstrate the value of alternative behaviors and management systems.

Exhibit 22 shows how the experiential participative model works, especially its strong similarities to the non-experiential participative model.

a. Inputs. There are three inputs to the model: police managers unfamiliar with participative theories and practices, instructors well versed in these theories and practices, and a non-stressful setting.

b. Processes. Ten major in-program processes occur. Trainees are presented with participative management theories and their interrelationships. MBO is described to them and analyzed as an operationalized form of participative management. They take self-assessment exercises to determine individual attitudes and orientations on key management dimensions, followed by a debriefing on assessment results. MBO is compared with current police management practice, and the advantages and disadvantages of MBO in relation to it are weighed. Trainees are presented with a description and analysis of the problems, preconditions, and precedents for the implementation of MBO. They take case study exercises that elicit trainee management philosophy and allow its shaping toward participative principles through application to hypothetical situations. They also perform active individualized exercises, both written and experiential, including those on setting clear and measurable objectives. Complementing the individualized exercises, they participate in competitive and structured group experiences and simulation exercises. After each exercise, instructors debrief trainees to illustrate relevant management principles and to help analyze individual and group responses.

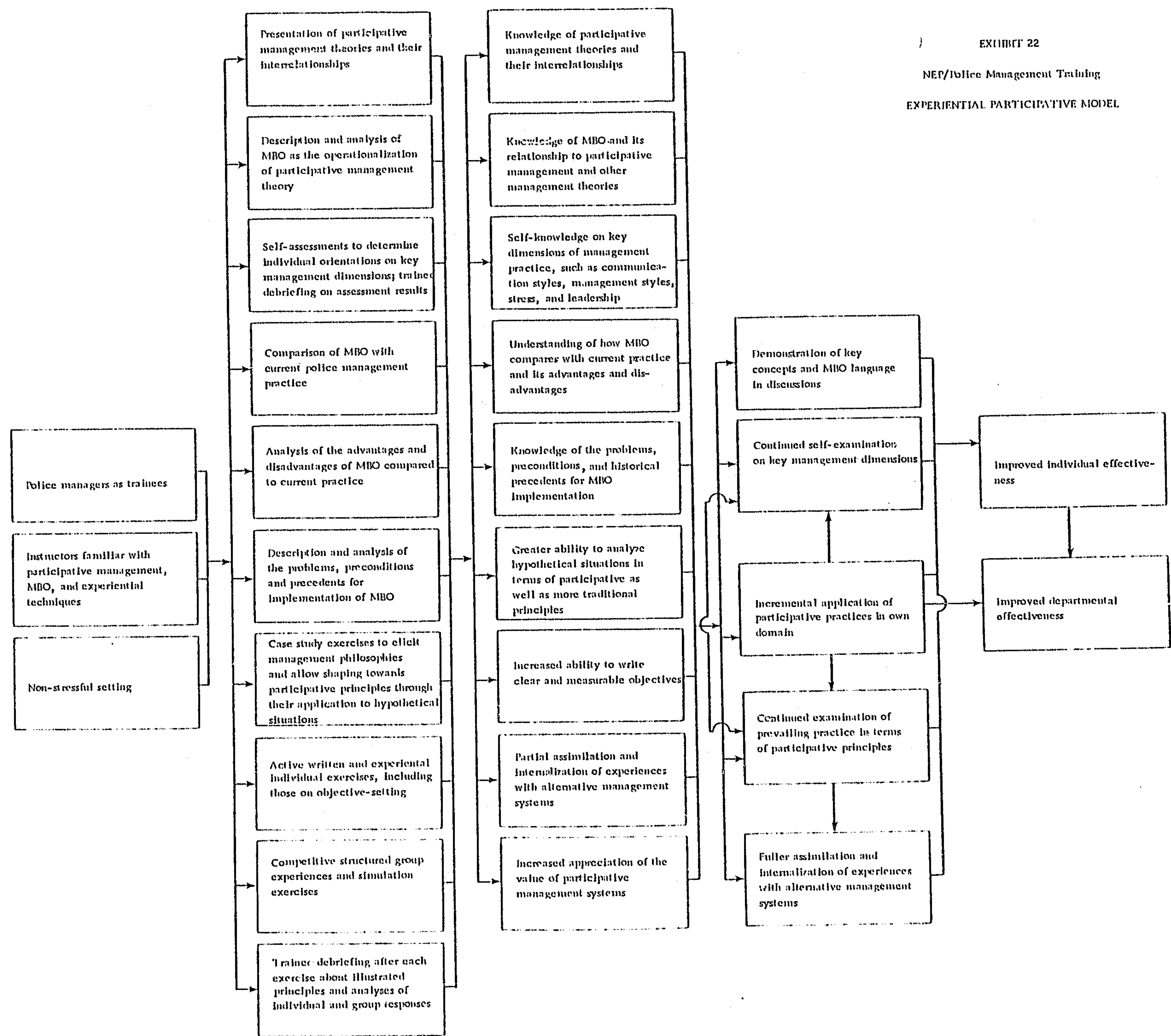
c. In-program outcomes. Nine in-program outcomes may be expected. Trainees leave the program with knowledge of participative management theories and their interrelationships. They are familiar with MBO and its relationship to participative and other management theories. They also know more about their own attitudes and orientations on key aspects of management behavior. They understand how MBO compares with current practice and also MBO's advantages and disadvantages in relation to it. They know the problems, preconditions, and precedents for MBO's successful implementation. They are better able to analyze management situations in terms of participative and other management principles and to develop clear and measurable objectives. Trainees leave the program having partially assimilated and internalized these experiences with alternative management systems and with an increased appreciation of the value of participative management.

d. Impacts. The in-program outcomes produce five short-term impacts. Trainees demonstrate key participative concepts and MBO language in post-program discussions. They continue to examine their attitudes and orientations on management behavior. This self-examination is accompanied and supported by the incremental

EXHIBIT 22

NEP/Police Management Training

EXPERIENTIAL PARTICIPATIVE MODEL



implementation of participative practices in trainee departments. Trainee self-examination and incremental application of recommended practices support continued trainee examination of prevailing management practice in terms of participative principles. Through these post-program behaviors, trainees more fully assimilate and internalize their experiences with alternative management systems. The ultimate expected impacts are improved individual and departmental effectiveness.

D. AUXILIARY MODELS

The auxiliary training models focus on training's anticipated effects on the work place or on the larger system environment. Their processes occur in parallel to the processes of the basic models or even after the completion of training. This section contains six auxiliary models: "greasing the skids," certification, network, anointing, departmental decision-making, and critical mass models. These auxiliary models can be linked to any prescriptive or participative model or combination of models. They can also be linked to each other, although the compatibility among them varies.

1. "Greasing the Skids" Model. The "greasing the skids" auxiliary model emphasizes that most programs provide substantial amenities to trainees, inside and outside the classroom, to ensure that trainees return to their jobs rested, uplifted, and satisfied with the training environment. It looks upon these amenities and the informal interaction among officers from several departments as means to boost morale and performance, independent of instructional contents. Consistent with this model, police executives often decide to which programs officers should be sent based heavily on consideration of these ancillary effects, regardless of what information trainees might learn from a program. The "greasing the skids" auxiliary model is based on the following underlying assumptions, as articulated by program administrators, police executives, training directors of departments, and trainees themselves:

- Police executives need a mechanism, short of promotions and pay increases, by which to reward past performance, improve present performance, and boost individual and departmental morale.

- Training programs often offer amenities geared to maximize trainee satisfaction with the program, ultimate acceptance of its contents, and repeat business.

- The improved morale and performance of recently trained officers can boost morale, show that performance will be recognized and rewarded, and provide a model for others.

- Training programs can provide executives with a mechanism to reward individual performance, boost morale, and stimulate improved group performance.

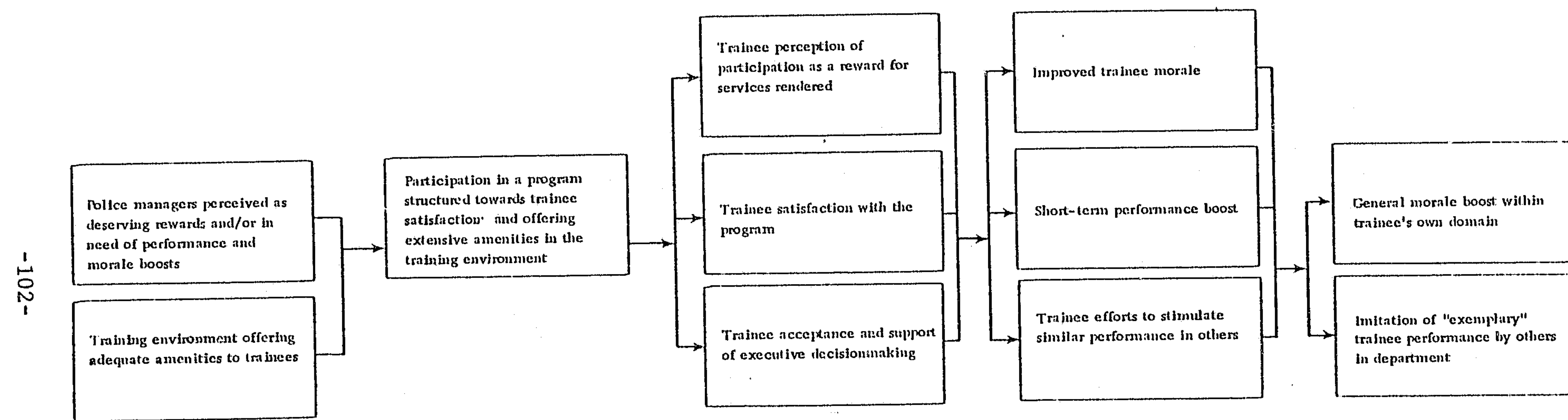
Exhibit 23 shows how the "greasing the skids" model works.

a. Inputs. The model specifies two inputs: police managers seen as deserving rewards and/or in need of performance and morale boost and a training environment with adequate amenities.

EXHIBIT 23

NEP/Police Management Training

"GREASING THE SKIDS" AUXILIARY MODEL



INPUTS | PROCESSES | IN-PROGRAM OUTCOMES | SHORT-TERM IMPACTS | LONG-TERM IMPACTS

b. Processes. The only process specified by this model is trainee participation in a program that offers discernible and adequate amenities and is structured to satisfy social and recreational needs.

c. In-program outcomes. Participation leads to three in-program outcomes: trainees are satisfied with the program, perceive their participation as a reward for past performance, and more strongly accept and support executive decision-making.

d. Impacts. As results, when they return to their jobs, trainees display short-term boosts in morale and overall job performance and attempt to stimulate similar performance by other officers. This leads to better departmental morale and the imitation of trainee "exemplary" performance by other officers.

2. Certification Model. The certification auxiliary model shows how a mandated and effectively enforced certification program can support immediate program outcomes, perpetuate them, and further supplement them by impacts on the overall system. It supports immediate program outcomes by enforcing minimum standards of trainee course performance and by awarding certificates, credentials, and other incentives to officers who complete the program. It perpetuates program outcomes by the promotion of increased career development, additional education and training, and the reinforcement of previous programs in a certification sequence. It further supplements these outcomes by certain systemic impacts, such as the attraction of more qualified individuals to law enforcement careers, reduction of departmental turnover, improved reputation for law enforcement, and increased allotment of resources to law enforcement. The certification auxiliary model is based on the following underlying assumptions articulated by instructors and program administrators and oversight agency representatives in certification programs:

- The development of police management capability has been hindered by the unequal abilities among police managers, the lack of an agreed-upon body of essential functional information, the absence of enforced standards for minimum job knowledge and acceptable performance, and a resultant poor public reputation for law enforcement managers.

- Lack of certification and credentialing processes to recognize police managers as professional managers--with skills and knowledge far beyond minimum requirements for police work--further inhibits development of police management capabilities.

- A certification program can improve receptivity to training, ensure that managers are well versed on current methods of supervision and management, and provide motivation to pursue further personal and career development.

- A certification program can improve public perception of police managers by fostering a professional image.

- By enforcing minimum knowledge standards, a certification program can effectively weed out managers of marginal competence.

- By fostering a professional police image and encouraging career development, certification and credentialing processes can attract more qualified applicants while decreasing turnover.

Exhibit 24 shows how the certification model works.

a. Inputs. There are three inputs: police managers from the certifying jurisdiction, program staff, and state certification requirements.

b. Processes. The only process that this model can specify is participation in a program operated in accord with state certification requirements.

c. In-program outcomes. Six in-program outcomes may be expected. Most trainees attain the minimum knowledge and performance levels required for certification. They receive credentials that recognize the knowledge attained as well as past managerial experience. They obtain financial or material incentives for successful program completion. They also leave the program with an increased sense of police professionalism and an increased interest in career development, further education, and additional training. Simultaneously, the enforcement of certification requirements weeds out managers of marginal or substandard competence.

d. Impacts. Five short-term impacts follow. Trainees improve their individual performance. They undergo increased career development, further education, and additional training. More qualified individuals are attracted to careers in law enforcement. Departmental turnover is reduced. Previous offerings of a certification program sequence are reinforced by more acceptable job behavior. As further results, departmental effectiveness and the statewide reputation of law enforcement are improved, increased resources are allotted to law enforcement, and statewide law enforcement capabilities are increased.

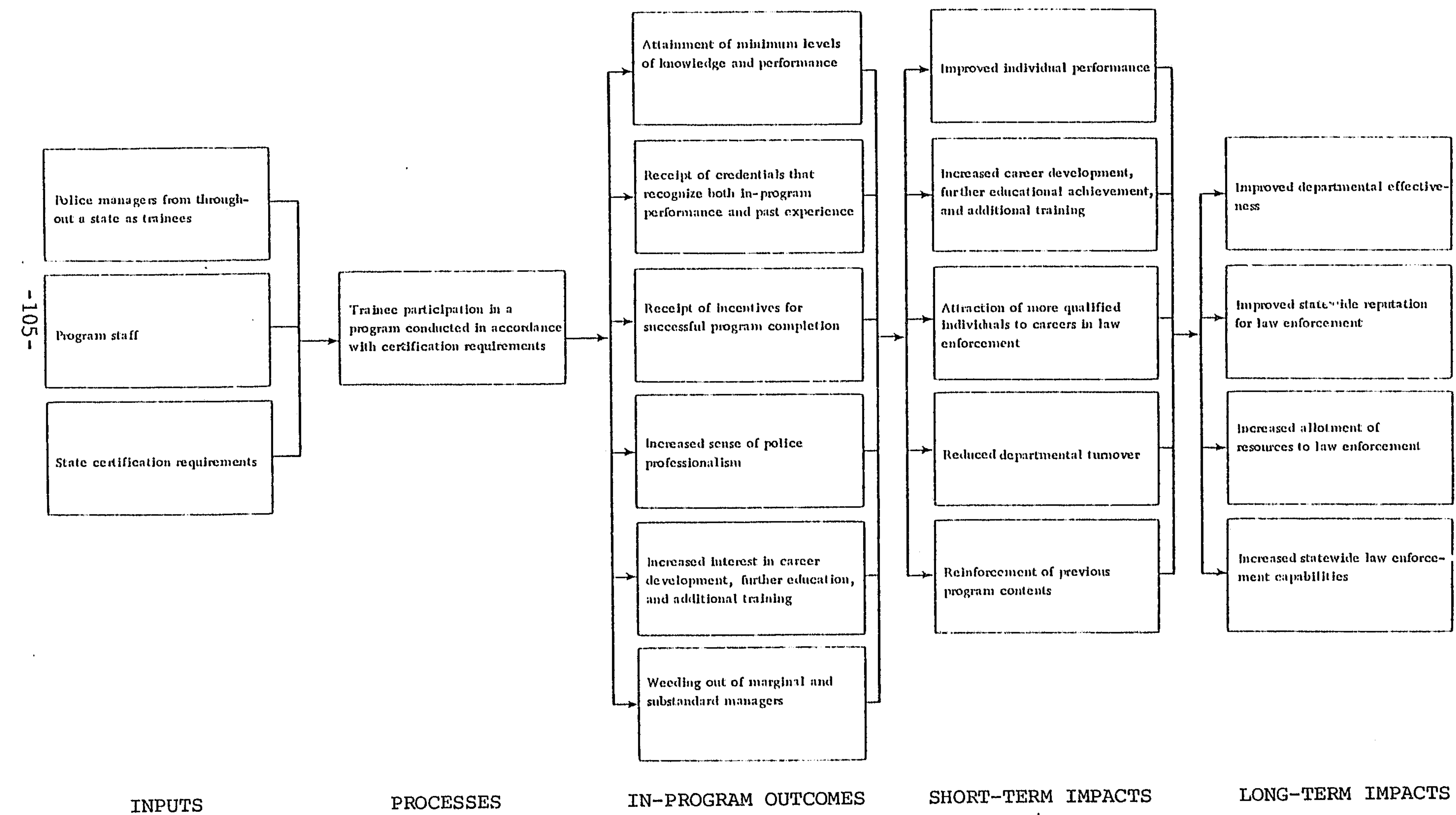
3. Network Model. The network auxiliary model uses informal trainee social interactions during a course to extend and perpetuate course effects beyond a course's conclusion. This model can flourish only in residential training programs that bring together managers of diverse backgrounds for an extended duration, either in a single long course or through a sequence of short courses. It flourishes only in such programs because they allow prolonged and continual interaction of trainees during and between formal training sessions and also during off-time hours. This intense interaction leads to lasting personal acquaintances, which form the basis for post-program network activities. Such interaction is inhibited by non-residential or short-term residential programs.

The assumptions tied to the network model limit its occurrence to prescriptive and participative basic program models. The formation of a network within a program involves the systematization of already existing informal processes among trainees and, thus, includes both formal and informal in-program activities. These activities provide the stimulus for formation of the network, which initiates its own activities after program completion. The purposes for network activities coincide with program goals and promote their realization through six distinct processes: reinforcement of immediate program outcomes, perpetuation of program outcomes through network activities, assistance to graduates in

EXHIBIT 24

NEP/Police Management Training

CERTIFICATION AUXILIARY MODEL



management problem solving, promotion of organizational change, assistance to graduates in career development, and promotion of personal fulfillment in graduates. The network auxiliary model is based on the following underlying assumptions, as articulated by trainees and program managers who had experienced success in the development and maintenance of a network:

- To anticipate long-term impacts realistically, training programs need a mechanism for extending their influence beyond a course's duration
- Especially in long-term residential programs, informal interaction processes can promote the formation of lasting acquaintances.
- A network or fraternity of program graduates can help programs to extend their intervention and allow program graduates to benefit from the experiences of other graduates.
- A network can contribute to continued program maintenance and growth, while fostering realization of program goals and objectives.
- Network activities can increase perceptions of police professionalism among graduates and further their personal and career development.
- Network activities can also promote effective adoption of major organizational change.

Exhibit 25 shows how the network model works.

a. Inputs. There are three specified inputs: police managers with varied backgrounds and experience from several departments, a program staff that is on location and sensitive to the possibilities for using informal trainee interaction to realize program goals, and a residential facility.

b. Processes. Ten major processes occur during a course. Trainees interact formally through in-class introductions, structured classroom interaction, and participation in competitive group activities. They interact informally in off-time group projects and in recreation. Instructors support these processes through carefully planned seating and rooming arrangements and control over informal interaction in scheduled social activities. Parallel to these activities, three other processes occur that directly reinforce the network's value. Program graduates make presentations and deliver a valedictory address. Graduates and instructors explain formal opportunities to perpetuate the interaction developed in the program and exhort trainees to continue to rely on each other and to call on program staff freely.

c. In-program outcomes. Nine in-program outcomes are expected to follow. Trainees are satisfied with the program and accept its contents. They appreciate the personal value of acquaintances made during the course and better recognize the management problems that they share with other officers. They leave the program with increased group cohesiveness, reduced insularity, and an improved sense of police professionalism. They recognize other graduates as professional resources and look upon program staff both as an informational resource and as a liaison for contacts with other graduates. As a result, trainees leave the course formally initiated into an existing program network.

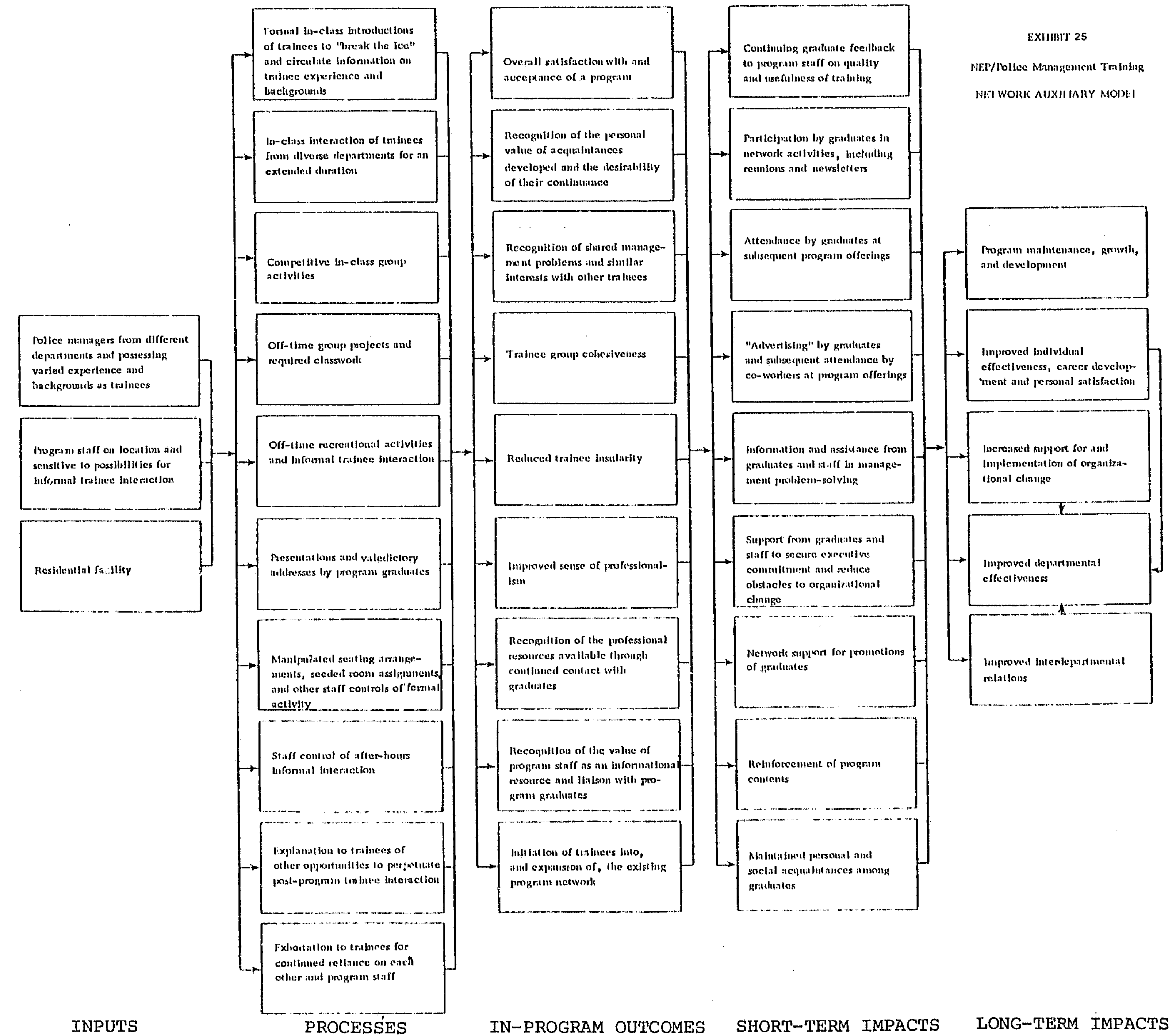


EXHIBIT 25
NEP/Police Management Training
NETWORK AUXILIARY MODEL
Police managers from different departments and possessing varied experience and backgrounds as trainees
Program staff on location and sensitive to possibilities for informal trainee interaction
Residential facility
Formal in-class introductions of trainees to "break the ice" and circulate information on trainee experience and backgrounds
In-class interaction of trainees from diverse departments for an extended duration
Competitive in-class group activities
Off-time group projects and required classwork
Off-time recreational activities and informal trainee interaction
Presentations and valedictory addresses by program graduates
Manipulated seating arrangements, seeded room assignments, and other staff controls of formal activity
Staff control of after-hours informal interaction
Explanation to trainees of other opportunities to perpetuate post-program trainee interaction
Exhortation to trainees for continued reliance on each other and program staff
Overall satisfaction with and acceptance of a program
Recognition of the personal value of acquaintances developed and the desirability of their continuance
Recognition of shared management problems and similar interests with other trainees
Trainee group cohesiveness
Reduced trainee insularity
Improved sense of professionalism
Recognition of the professional resources available through continued contact with graduates
Recognition of the value of program staff as an informational resource and liaison with program graduates
Initiation of trainees into, and expansion of, the existing program network
Continuing graduate feedback to program staff on quality and usefulness of training
Participation by graduates in network activities, including reunions and newsletters
Attendance by graduates at subsequent program offerings
"Advertising" by graduates and subsequent attendance by co-workers at program offerings
Information and assistance from graduates and staff in management problem-solving
Support from graduates and staff to secure executive commitment and reduce obstacles to organizational change
Network support for promotions of graduates
Reinforcement of program contents
Maintained personal and social acquaintances among graduates
Program maintenance, growth, and development
Improved individual effectiveness, career development and personal satisfaction
Increased support for and implementation of organizational change
Improved departmental effectiveness
Improved interdepartmental relations
INPUTS
PROCESSES
IN-PROGRAM OUTCOMES
SHORT-TERM IMPACTS
LONG-TERM IMPACTS

d. Impacts. Nine short-term impacts flow from these outcomes. Trainees provide continuing feedback to program staff on the quality and usefulness of training. They participate in network activities, such as reunions and newsletters, as well as in subsequent program offerings. They also "advertise" the program to co-workers, to encourage their attendance. Program staff and the network provide graduates with information and assistance in management problem solving, career development, and securing of executive commitment to change. These network activities reinforce program contents while maintaining acquaintance among graduates.

From these short-term impacts, five long-term impacts are expected to follow. The program is maintained and continues to grow and develop. Trainees experience improved individual effectiveness, career development, and personal satisfaction beyond that obtained from the formal program. There is increased support for and implementation of organizational change in agencies affected by the network. Relations among trainee departments are more cooperative. As a result, affected departments are more effective.

4. Anointing Model. The anointing model demonstrates the use of training to recognize managers already tagged for promotion to senior positions. It expects attendance at certain nationally recognized programs to enhance career progression and stature through esteemed new credentials and influential new acquaintances. At the same time, it expects the nationally recognized programs to maintain their reputations through the career development and heightened stature of their graduates, who have been aided by program attendance. This model resembles the network and critical mass models in its emphasis on the career development of graduates. It is based on the following underlying assumptions, as articulated by staff from major national programs and by executives from major departments utilizing these programs to recognize promising managers:

- To be effective, senior police officals must possess stature and credibility based on outstanding past performance and training: an "anointing" that distinguishes them from other departmental personnel.
- The stature of senior police officials can be increased through amicable relationships with senior officials from other departments.
- The visibility and reputation of major national providers of police management training are closely tied to the senior positions held by graduates.
- Because such programs are expensive to operate, attendance can be open only to a select group of highly promising officers.

Exhibit 26 shows how the anointing model works.

a. Inputs. The anointing model specifies two inputs: managers selected for immediate or eventual promotion to senior positions and nationally recognized police management training programs.

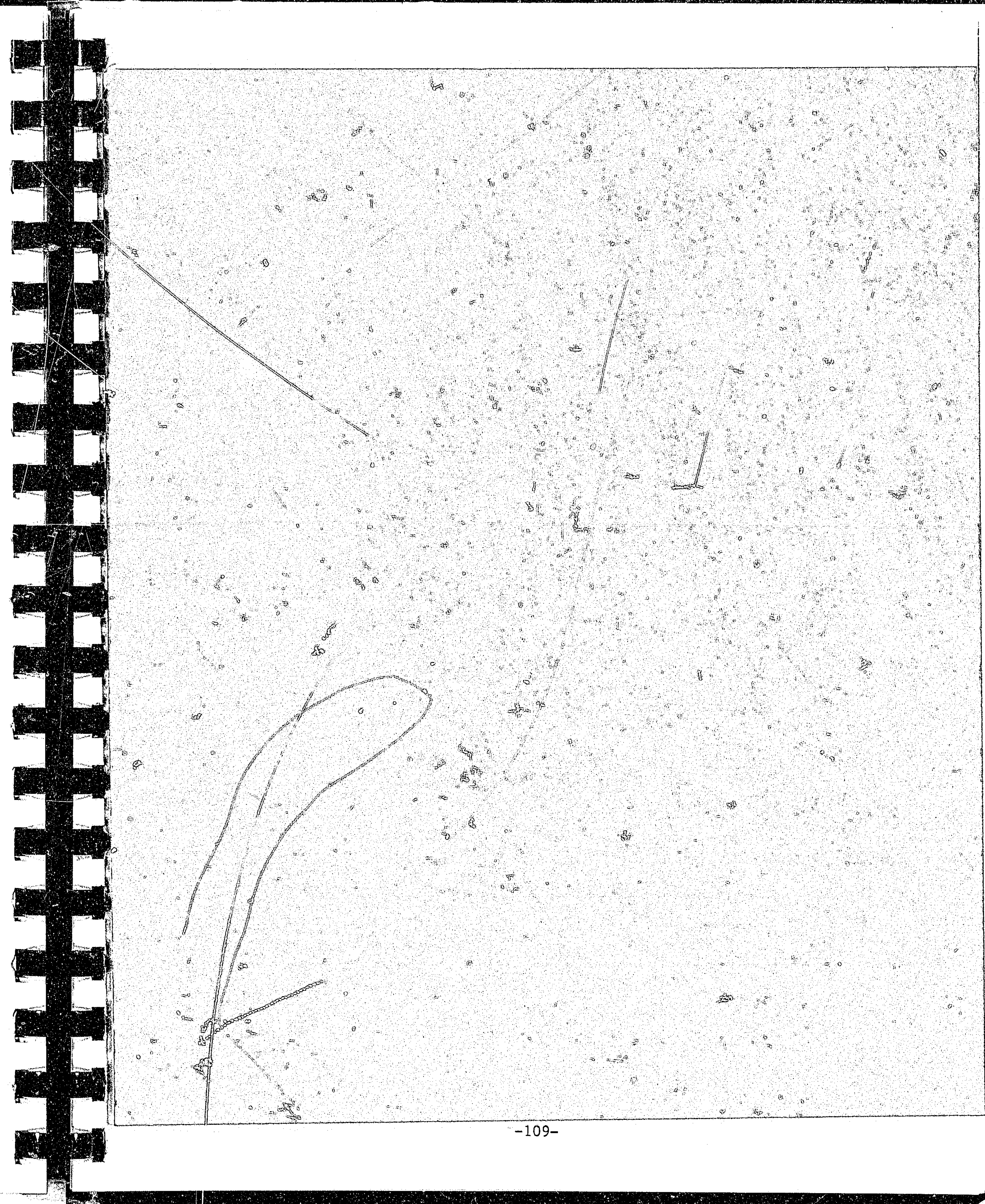
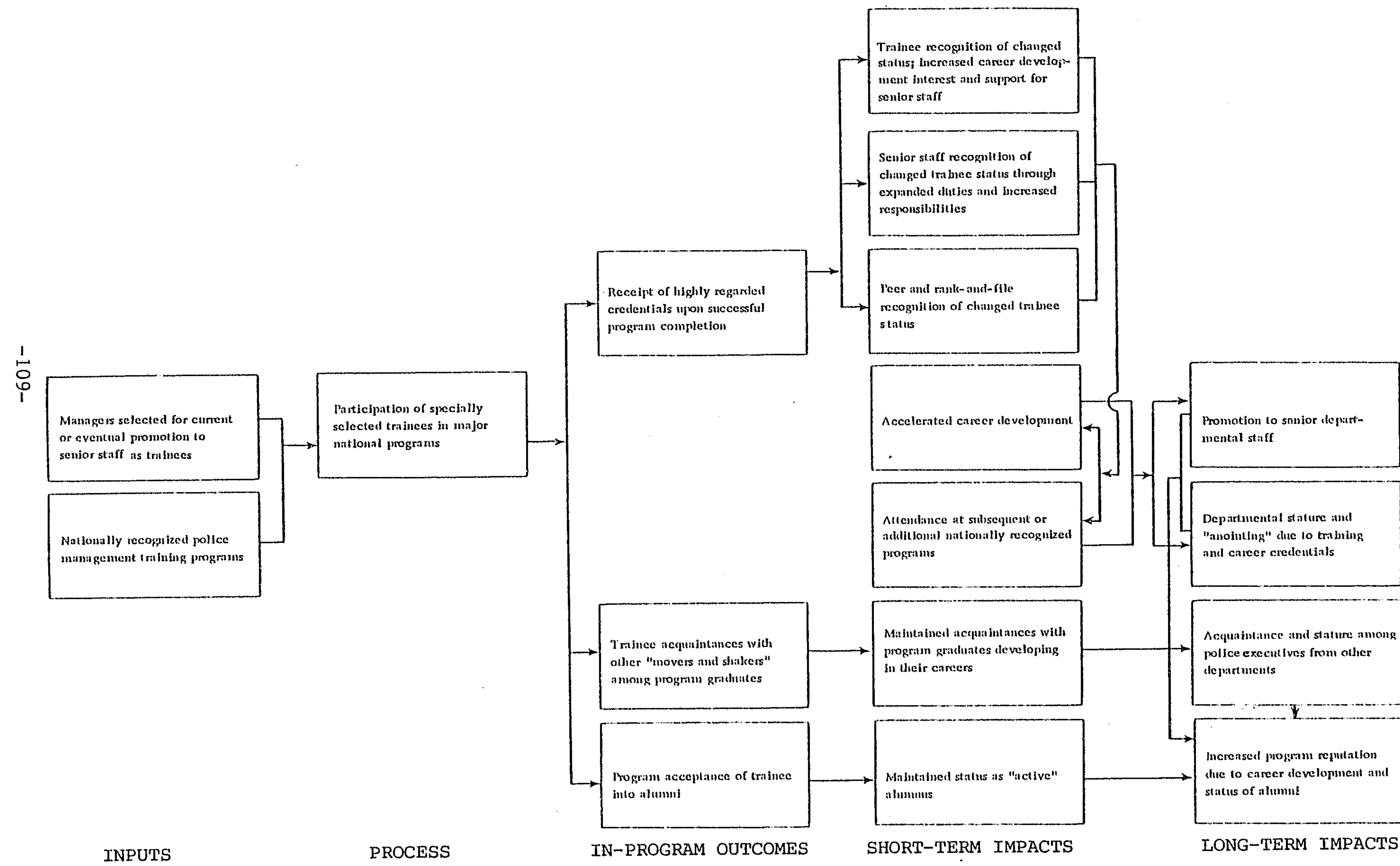


EXHIBIT 26

NEP/Police Management Training

ANOINTING AUXILIARY MODEL



b. Processes. The process that triggers the anointing function is attendance by selected managers at nationally recognized programs. This process cannot be further specified.

c. In-program outcomes. Three outcomes follow from trainee attendance at major national programs. Trainees receive highly regarded credentials for successful program completion. They make acquaintances with other trainees who are "movers and shakers" within their own departments. They are accepted by the program as alumni.

d. Impacts. Seven short-term impacts are expected to follow. Trainees recognize their changed status and, consequently, show increased orientation to career development and support for senior department staff. Senior staff, in turn, respond to changed trainee status with expansion of their duties and responsibilities. Peers and the rank and file also recognize changes in status. As a result, trainees undergo accelerated career development and attend other nationally recognized programs. At the same time, graduates maintain acquaintances with other graduates, who are undergoing similarly accelerated career development and continue to support the program as active alumni. These short-term impacts produce four long-term impacts. Trainees eventually obtain senior staff positions. In those positions, their stature is enhanced through their recognized training and career credentials. Their stature is also enhanced by continued acquaintance with graduates from other departments, who have by now achieved senior staff positions. At the same time, program reputation is maintained and increased through the career development and status of alumni.

5. Departmental Decision-making Model. The departmental decision-making auxiliary model uses training as a communications vehicle between senior departmental staff and line managers. It differs from the compliance models in that communication flows two ways, whereas, in the compliance models, the flow is only downward. It provides senior departmental staff with critical information about the feasibility of a decision and its likelihood for successful implementation. It provides line managers with information about a pending decision and with an opportunity to shape the decision through critical feedback. This two-way information flow is seen as assuring that the best possible decision will be made and that implementation efforts will be smooth. The departmental decision-making auxiliary model is based on the following underlying assumptions, as outlined by departmental officials and instructional staff:

- Senior police officials need complete and accurate information about the feasibility of departmental decisions and appropriate means for implementing them.

- The implementation of departmental decisions can fail for any of three reasons: senior officials do not obtain adequate feedback, the rank-and-file perceive that they have not been a party to the decision, and information about implementation is distorted through inadequate communication.

- Training programs can serve as conduits for valuable rank-and-file feedback to senior officials in making decisions and implementing changes.

- Training programs can also provide trainees with a sense of having contributed to a decision, thus helping to secure their commitment toward implementation of the decision.

Exhibit 27 shows how the departmental decision-making model works.

a. Inputs. There are three inputs: managers from departments contemplating a major decision, senior departmental officials as instructors, and training staff.

b. Processes. Seven major processes occur. Trainees are presented with the details of the management problem to be addressed by a prospective decision. Management options involved in the decision are described and analyzed. Senior departmental officials present their management philosophies and approaches to the decision. The results of the prospective decision are compared with current practice, and the advantages and disadvantages of each are weighed. Trainees are presented with the preconditions, problems, and precedents for implementing the decision based on experiences of other departments. They are provided with an opportunity to make an informal critique of the prospective decisions and to contribute related information to senior officials.

c. In-program outcomes. Six in-program outcomes are expected to follow. Trainees know about a prospective decision and the problem it addresses. They are familiar with the advantages and disadvantages of the decision. They know the preconditions, problems, and precedents for implementing the decision that have been experienced by other departments. As a result, they possess the capability to critique the prospective decision. They perceive that they have participated in departmental decision-making. Thus, a feeling of group consensus about the feasibility of the decision emerges, and trainees are better prepared to accept the decision's implications.

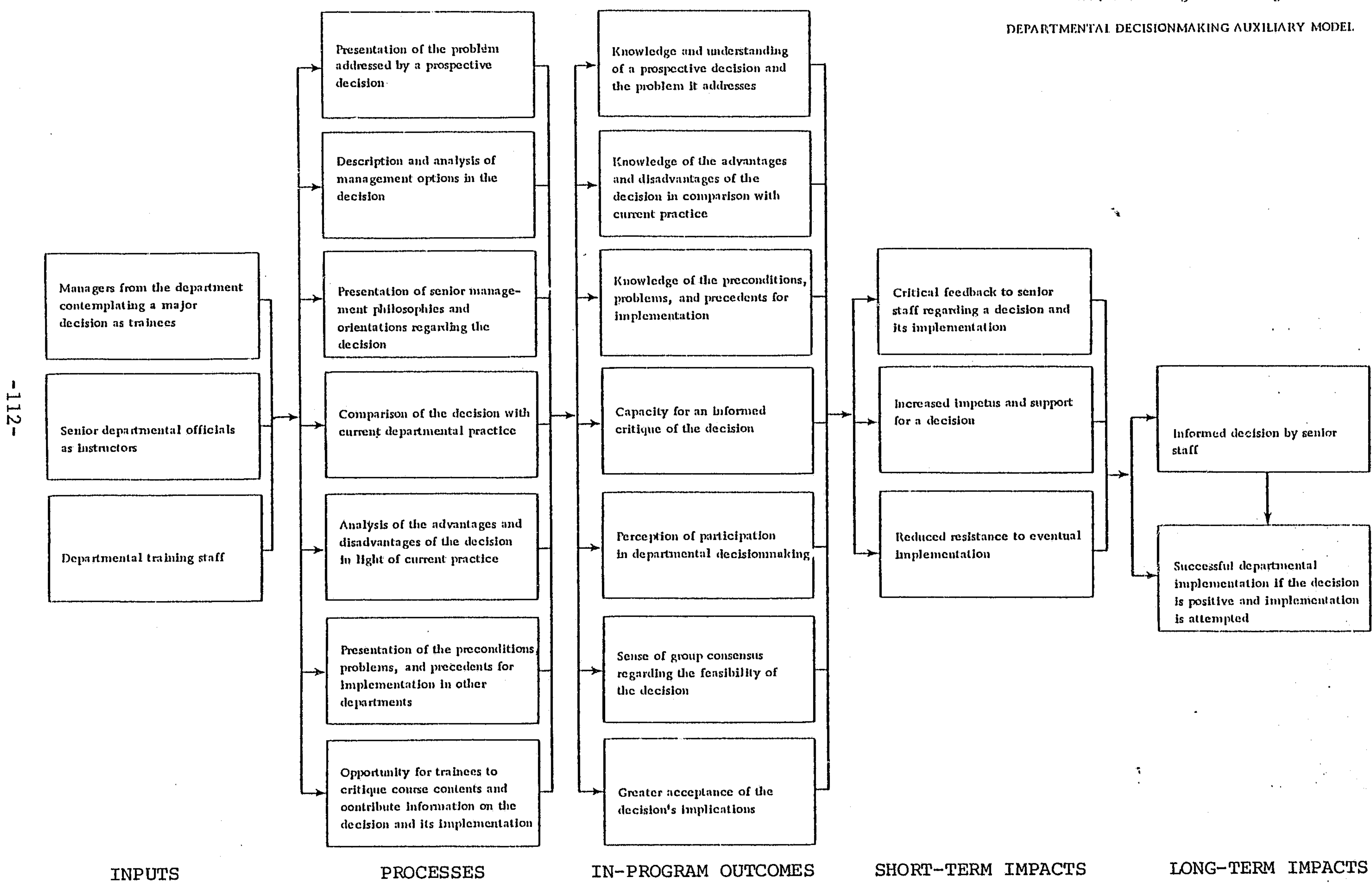
d. Impacts. As a result, three short-term impacts occur. Senior staff obtain critical feedback regarding a decision and its implementation. Increased departmental support for the decision is generated. Resistance to implementation of the decision is reduced. Ultimately, senior staff make an informed decision, and, if this decision is positive, implementation is smooth.

6. Critical Mass Model. The critical mass auxiliary model recognizes that training can stimulate organizational change but that change is contingent upon factors not under program control. For effective organizational change, there must be enough similarly attuned managers to carry out change and enough senior officials who support change and can initiate and oversee its implementation. The critical mass model involves activities geared to create the preconditions for change. Program staff encourage graduates to advertise the program and its concepts. Graduates assist other graduates in career development and incrementally implementing program concepts. These activities lead to the development of a critical mass: a sufficient number of graduates throughout a department and a sufficient number of graduates in senior staff positions as proponents of change. Once the critical mass is reached, change can follow. This model is related to the network model in its reliance on post-program relationships and interactions among graduates. It is based on the following underlying assumptions, as articulated by program administrators and instructors in programs espousing a critical mass concept.

EXHIBIT 27

NEP/Police Management Training

DEPARTMENTAL DECISIONMAKING AUXILIARY MODEL



- Because police agencies generally resist attempts at innovation, organizational change is typically a long-term incremental process.
- A key resource in organizational change is the presence of a sufficient number of trained managers throughout a department to carry out implementation effectively.
- Another key resource is the presence of trained senior staff who are capable of initiating and supervising change.
- Training can more realistically expect to promote organizational change if it can mobilize a critical mass, consisting of these two resources
- Program graduates in positions of authority can help produce this critical mass by assisting in the career development of later graduates while promoting continued participation in the program by other graduates and increased participation by co-workers.

Exhibit 28 shows how the critical mass model works.

a. Inputs. The model specifies two inputs: management trainees and program staff oriented toward organizational change.

b. Processes. The only specifiable in-program process is trainee participation in activities promoting large-scale organizational change.

c. In-program outcomes. The in-program outcomes are an increased pool of program graduates and trainee satisfaction with and commitment to the program.

d. Impacts. Six short-term impacts result. Graduates "advertise" the program to co-workers, who attend subsequent sessions. They assist other graduates in promotions and career development. They undergo accelerated career development themselves. They incrementally implement changes recommended by the program in their own domains. They also assist other graduates in the incremental implementation of change. As a result, departmental acceptance of and support for change increases. These short-term impacts lead to five long-term impacts. A sufficient number of program graduates for the implementation of organizational change develops. A sufficient number of graduates reach the positions of authority necessary to initiate and support organizational change. An organizational foundation is laid for the acceptance of large-scale change. A critical mass of graduates with sufficient numbers and authority to implement large-scale organizational change successfully thus emerges. This critical mass leads to successful implementation of large-scale organizational change.

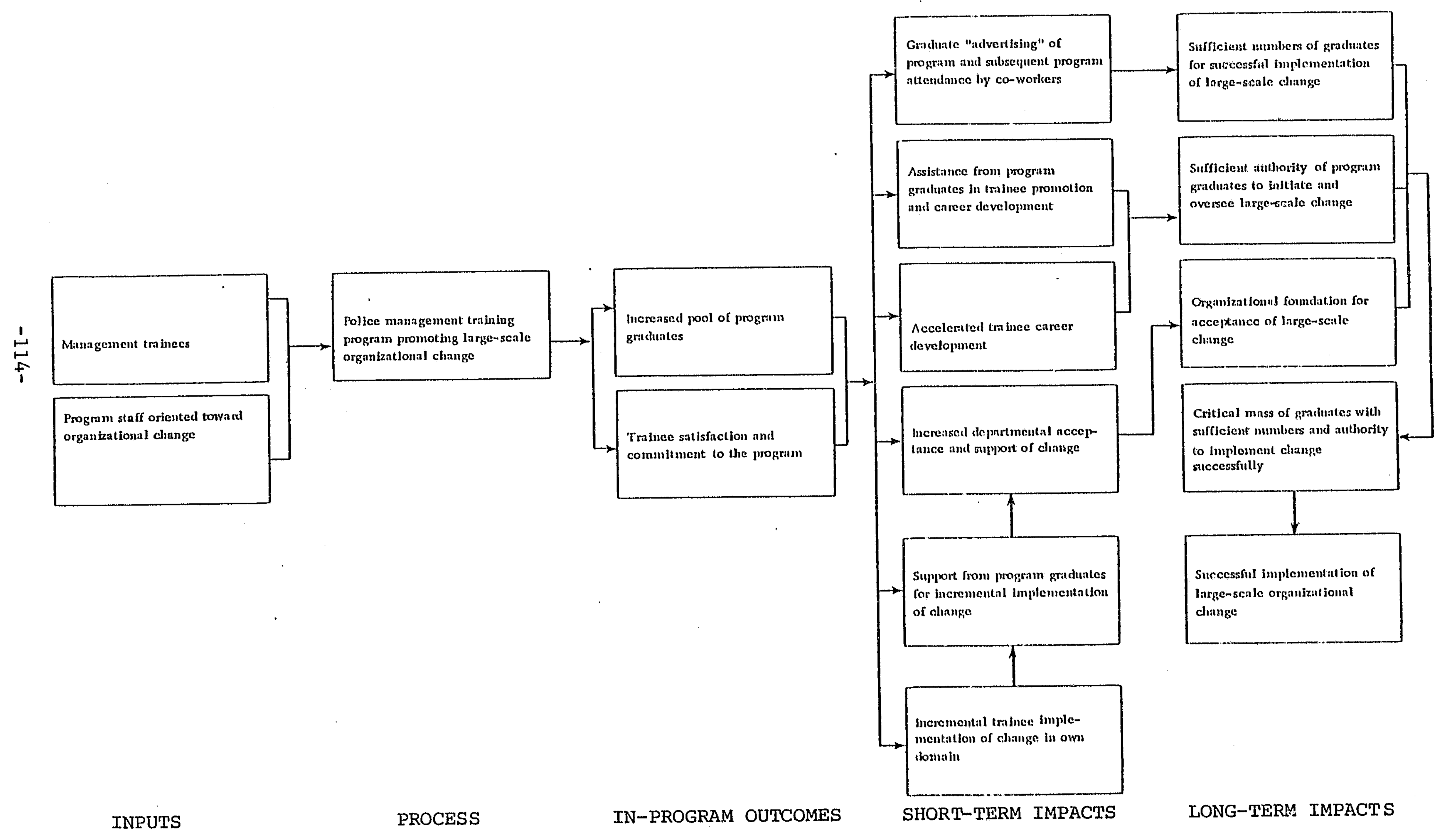
E. MIXED MODELS

No single model was either fully articulated or unequivocally espoused by the programs we observed. In other words, none of the programs fully expressed any model and all mixed several models together. In saying that no program fully articulated any model, we mean that, to construct models for the programs that we observed, we had to tease out and piece together bits of information

EXHIBIT 28

NEP/Police Management Training

CRITICAL MASS AUXILIARY MODEL



gained in observation, interviews, and review of program documentation and files and from other sources. We abstracted the empirical models in this chapter from the observed program models, but the observed programs did not elaborate them in this form or level of detail. In saying that all programs mixed models, we mean that, though a single basic model often predominated, each program mixed two or more of the eight basic models. With the exception of some programs strongly oriented toward compliance, each program also appended one or more auxiliary models to the mix of basic models, to show the larger impacts sought on a department or system level. (It is intrinsic to the auxiliary models, in fact, that they cannot stand on their own but must be tied to basic models to find their substantive content.)

The nature and causes of this model mixing vary among and even within programs. Much of it is officially recognized, is set forth in public descriptions of programs, and fits together comfortably. But a lot of mixing also stems from lack of coordination within programs of curricula and instructional personnel and results in an unintended "smorgasbord" type of program that points trainees simultaneously in several directions and, ultimately, in no clear direction at all. A lot of model mixing also stems from the different roles and responsibilities that people have in a program.

How an individual or group views the operations and goals of a program can be directly related to their responsibilties for program development. In presenting each model above, we noted the sources most frequently expressing its concepts. We can illustrate the effects of vantage point well through examples. Those directly responsible for instruction usually stated the basic model(s) in which a program operated more clearly than did administrators, who tended to gloss over differences in content and to focus on global "preparation to fill the specific needs of the trainee's chief." Those in instructional roles were also more likely than administrators to extend the implications of their emphases in course content to the build-up of a critical mass prepared for organizational change. In contrast, administrators of major residential programs and their graduates were more inclined to describe their program in terms of the "network" of graduates being built up than were individual instructors, except where they too were graduates. Trainees, police executives, and departmental training directors tended to look at programs for their effects on motivation and "greasing the skids" and also for the "anointing" effects of training received from major national providers. The administrators of statewide certification programs were most likely to highlight certain aspects of the certification model, such as weeding out inept managers and advancing police professionalism. Clearly, what one contributes to and stands to gain from a program influence the expectations one is likely to hold for it.

The problem is, this mix or "coexistence" among several models in a single program often produces ambiguity about the model or models in which the program is operating and, for that matter, those by which it ought to operate. As a result, people develop divergent notions about trainee selection, staff hiring, instruction coordination, compliance with state program requirements, program amenities, needs assessment procedures, curriculum design, and other matters. Thus, those with different functional responsibilities often do not act in concert.

Much of this model mixing is legitimate and useful. At the least, a lot of it is inevitable because people have different functional responsibilities in program development and also hold different management philosophies. Regardless, the variation among models and the phenomenon of model mixing have enormous implications for how a program should be managed and, as a corollary, how it should be evaluated. It is essential that a program's managers and operators determine carefully and exhaustively the models by which its key contributors actually operate. This is a precondition for effective program management and useful program evaluation.

Five: Program Evaluability

What types of obstacles impede both useful program evaluation and effective program management? What can be done to overcome them? Throughout this report, we have emphasized the central role that evaluability assessment can play in both evaluation design and program management. Evaluability assessment is really little more than an effort to define a program and to identify a useful role for evaluation in the program's management. In Chapter One and Appendix 2, we discussed how periodically answering certain preliminary questions not only leads to more useful evaluation designs but also helps identify roadblocks to effective program management and the documentation of a program's success. We said that, when a program does not meet the conditions for evaluability, its manager will encounter difficulties in making it work, demonstrating how well it works, or both.

We also outlined a broad set of potential roles for the person assigned evaluation responsibilities. The basic responsibility is to describe the actual program fully, including the roadblocks to evaluability. The evaluator's roles extend beyond the documentation of the program's shortcomings, however. A second role is to work with program managers and operators to identify ways to remove roadblocks to evaluability. A third is to help select a strategy to improve the program. A fourth is to monitor the implementation of the strategy. A fifth is to design and complete an evaluation approach that meets the program's needs. These evaluation activities primarily require common sense and an open mind. With the possible exception of the last evaluation responsibility, they do not require the evaluator to have special training or skills in traditional evaluation approaches.

The purpose of this chapter is to provide a framework for the conduct of an evaluability assessment. The most practical information in this chapter is in a self-assessment guide that we call the "Evaluability Checklist." The Checklist consists of seven tabular exhibits, each of which corresponds to one of the seven evaluability criteria. In relation to each of the seven criteria, we identify relevant roadblocks to evaluability. For each roadblock, we specify one or more ways to eliminate it or at least reduce its potency. The Checklist may be used in two general ways. For a particular program, it may be used to identify relevant roadblocks to evaluability and to develop approaches for making the program more evaluable. On a larger scale, it may be used to assess the state of program evaluability and to develop concerted strategies for improving evaluability. The second use is made possible because we note the frequency with which programs across the country confront each roadblock. Either way, the Checklist may be used to illustrate the type of analysis that needs to be done or to guide the step-by-step analysis.

A. SEVEN CRITERIA FOR PROGRAM EVALUABILITY

This chapter is organized into seven exhibits, as noted above. The exhibits center on the seven basic criteria for an evaluable program: well defined (Exhibit 29), acceptable (Exhibit 30), valid (Exhibit 31), plausible (Exhibit 32), feasible data system (Exhibit 33), reliable data system (Exhibit 34), and plausible intended uses of data (Exhibit 35). These exhibits correspond to the seven basic questions that must be asked to determine a program's evaluability:

- Does program management define with reasonable completeness what is expected to happen in and result from the program?
- How acceptable is management's intended program in terms of what policymakers expect?
- Does the program actually in place validly represent program management's expectations?
- Is it plausible that the program will accomplish its purposes?
- Are the program's intended means for demonstrating success feasible?
- Does the program's intended data system have provision for repeated observations?
- Are management's intended uses of evaluation evidence under its control?

The seven exhibits share a common format. Each contains six columns: subquestions, roadblocks to evaluability, frequency, adjustments in expectations, adjustments in activities, and adjustments in information systems. The subquestions refine the basic question addressed by an exhibit into specific key issues. The roadblocks to evaluability show the practices and conditions that obstruct evaluability in relation to each subquestion. Frequency contains our estimate of the percentage of programs (expressed as High, Medium, or Low) that confront a given roadblock. It has no implications, however, for the relative importance or disabling effect of each roadblock. The three types of potential adjustments--in expectations, activities, and information systems--express alternate ways in which the program might be changed to mitigate a roadblock's disabling effects. Expectations refer to people's beliefs about what the program does and intends to accomplish. Activities refer to how the program selects and processes trainees in training delivery. Information systems refer to ways of collecting and ordering information to answer questions about the program's services.

A few words need to be said about repetition and use of language in the Checklist. There is necessarily some repetition or similarity within the roadblocks column and the three adjustments columns. The repetition or similarity in the roadblocks occurs because several roadblocks have implications for different aspects of evaluability and, hence, call for differing adjustments. The repetition or similarity in adjustments occurs because different obstacles to evaluation can sometimes be dealt with in analogous ways. Certain terms are used in the checklist that have not been defined. These include the terms

EXHIBIT 29(1)

NEP/Police Management Training

IS THE PROGRAM WELL DEFINED?

Basic Question: Does program management define with reasonable completeness what is expected to happen in and result from the program?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does the stated sequence of events leading to objectives begin with resources or activities over which the program has control?	Selection of trainees is largely in the hands of individual police agencies, not the training organization. This is especially true for field programs, because they typically do not come on site until training delivery.	M	Convey to participating agencies that selection of trainees who do not need training can be counter-productive.	Curtail services to agencies that persistently abuse selection criteria.	Develop profiles of each police agency's trainees to identify agencies that abuse selection criteria.
	The informal interactive processes inside and outside the classroom, although considered important to accomplishing the program's objectives, are beyond the program's control.	L	Unless informal activities can be structured and thus brought under control, downplay the role of informal processes in accomplishing objectives.	Manipulate seating arrangements, seed rooming assignments and discussion groups, structure evening recreational and group project activity.	Develop profiles of the background and training needs of each trainee prior to training, to provide the basis for structuring informal activity.
Does the sequence of events that management states must occur to achieve objectives encompass all key events in their intended order?	Management broadly defines the types of improved individual and agency performance it intends, with little or no attention to prior events. Management is unclear about what changes in knowledge, skills, or attitudes it expects to see by the end of a program session; what activities lead to these immediate outcomes; how these outcomes are to transfer to the job and affect individual performance; or how short-term transfer can produce larger departmental or system-wide changes.	H	Clarify expected causal links and key events by examining results of program monitoring, identifying points where expectations are unclear, and sorting out expectations into coherent chains of thought.	Modify actual activities to align better with clarified expectations.	Develop a system to monitor actual program activities and immediate outcomes to determine activities that might be plausible and/or in need of clarification.
	Management fluctuates and mixes concepts from several training models in its description of events that will produce its partially-defined objectives.	M	Clarify expected causal links and key events by sorting out expectations into coherent chains of thought, identifying points where diverse models reinforce or conflict with each other, prioritizing points of conflict, eliminating low priority conflicting expectations.	Modify actual activities to align better with clarified expectations.	

EXHIBIT 29(2)

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	Management does not recognize postprogram events outside the program's control needed to accomplish departmental or system-wide changes.	M	Take into account the extraprogram events that can intervene between short- and long-term effects of the program.		Develop a system of indicators for measuring impact and the environment's effects on it.
Does management anticipate and state the potential negative side effects of the program?	Management does not acknowledge that gross incompatibility between training-induced knowledge of innovative approaches and the trainee's current working conditions that are largely beyond his control can produce extreme frustration and potentially jeopardize career continuity.	L	Expect turnover to be high where implementation opportunities are low; take a longer-term view of the trainee's career options for implementation.	Prepare trainees for frustration; deal in training with problems in and preconditions for implementation.	
	Management does not acknowledge the potentially disruptive effects upon police agencies of indoctrinating trainees to management philosophies for which the agency is not prepared.	L	Recognize that organizational change is often a painful process, but that proper preparation can minimize disruption.	Deal in training with the process of preparing the agency for large-scale change.	
Does management's description of the program cover the key activities needed to produce each intended outcome and impact?	Activities are defined in terms of neither the immediate intended outcomes nor the instructional techniques needed to produce them, but mainly in terms of the general topics to be taught.	M	Clarify intended outcomes, the relationship of instructional methods and substance, and the relationship between activities and outcomes.	Eliminate activities that do not support intended outcomes.	Develop matrices for illustrating how topics, instructional techniques, and intended outcomes are linked.
	The intended interdependence among program activities is not defined, thus allowing random changes in the order of program components and modification of curricula without change in expectations.	M	Clarify how components are expected to build on each other, their necessary order, and how the intended order is tied to expectations.	Offer program components in an order that allows maximal achievement of objectives.	Develop system for measuring instructor and trainee views of how components interrelate and can best build on each other.
Does management specify the rationales for expecting activities to produce intended results?	Management does not specify the learning principles that link instructional techniques with desired results. This is especially important when dramatic changes in behavior or attitudes are expected to follow from traditional lecture and discussion techniques.	M	Clarify how immediate outcomes and later performance impacts relate to current and alternative instructional techniques.	Change instructional techniques so they better support intended results.	Develop matrices for illustrating how topics, instructional techniques, and intended outcomes are linked.
	Instructional rationales do not recognize the differences between teaching children and adults.	M	Define the need for using instructional techniques appropriate to an adult audience.	Implement techniques that encourage a more active learning role.	

EXHIBIT 29(3)

Subquestion	Roadblocks to Evaluability	Fre-quency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does management identify the evidence considered necessary to show that intended activities, outcomes, and impacts have occurred?	Program management does not define the measures and comparisons needed to show immediate program outcomes have occurred. This is especially important when the program aims at major changes in knowledge, attitudes, or behavior, but limits its success measures to rote paper-and-pencil exams and surveys of trainee reaction.	M	Highlight expectations for which measures and comparisons can be identified.	Reduce emphasis on activities for which alternate measurements and comparisons cannot be identified.	Identify measures and comparisons offering several levels of rigor in design and confidence in results, and adopt those that are necessary and sufficient to needs for evidence.
	Program management does not define the measures and comparisons needed to show that individual and departmental behavior have changed.	H	Highlight expectations for which measures and comparisons can be identified.	Reduce emphasis on activities tied to impacts for which alternate measurements and comparisons cannot be identified.	Identify measures and comparisons offering several levels of rigor in design and confidence in results, and adopt those that are necessary and sufficient to needs for evidence.
Does management define the range of values and time-frames within which resulting changes will be considered acceptable?	Management does not specify the extent to which change resulting from training is acceptable or desirable. This applies to any situation in which outcomes and impacts are described with phrases like "increased," "greater," or "movement toward." Failure to specify the desired range of values means that either too little or too much change might later be judged a failure.	H	Clarify the range within which change is considered acceptable, and outside of which change is considered unacceptable.	Coordinate trainer activities, modify instructional content, and adjust expectations held out to trainees to better accord with clarifications in acceptable range of values.	Make information systems sensitive to the range outside of which performance is unacceptable.
	Management does not specify the length of time needed for changes to take place. This is important in relation to expectations about when, and after what uncontrolled events have occurred, there ought to be demonstrable impact on trainee behavior and organizational performance. This is also important in relation to expectations about the amount of time needed for attitudes to measurably change: by the end of the program, or after return to the job and opportunity to test and personally validate new concepts.	M	Clarify the time-frames needed for change to take place.	Adjust expectations held out to trainees to better accord with clarifications in expected time-frames.	Develop information systems capable of measuring each specified point in time at which different observable changes ought to occur.
Does management specify the means for demonstrating that intended activities, outcomes, and impacts have occurred?	Management does not specify the information systems for measuring the program's outcomes, beyond paper-and-pencil examinations and surveys of trainee reaction, both inadequate to show the breadth or scope of intended results.	H	Highlight expectations for which measurement systems of acceptable costs can be identified.	Reduce emphasis on activities for which a measurement system of acceptable costs cannot be identified.	Identify alternate measurement systems and associated costs of operation, and select the one necessary and sufficient to meet needs for evidence.

EXHIBIT 29(4)

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	Management defines the measures and comparisons generally needed to show training has its intended effects on individual performance and agency effectiveness, but cannot translate these into measurement systems of known costs.	M	Highlight expectations for which measurement systems of acceptable costs can be identified.	Reduce emphasis on activities leading to impacts for which a measurement system of acceptable costs cannot be identified.	Identify alternate measurement systems and associated costs of operation, and select the one necessary and sufficient to meet needs for evidence.
Does management describe the actions to be taken on the basis of evaluation evidence and the ways these actions are expected to affect the program?	Program management defines actions to be taken based on evaluation evidence in terms of modifications in program components, but not in terms of major shifts in program thrust or potential termination of unsuccessful programs.	M	Outline the alternate uses of evaluation evidence for different users, and the levels of evidence needed to take different courses of action.		Develop matrices showing the users of evaluation evidence courses of action open to them, levels of evidence presumably needed to take action, action taken, actual levels of evidence.

EXHIBIT 30

NEP/Police Management Training

IS THE PROGRAM ACCEPTABLE?

Basic Question: How acceptable is management's intended program in terms of what policymakers expect?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Do policymakers and program managers agree on what resources the program ought to use and on what activities ought to take place?	Policymakers expect activities, especially, related to functional skills, that program management leaves out or at least de-emphasizes.	L	Bring policymakers and program management's expectations into alignment, to reflect changes in or maintenance of activities.	Increase emphasis on functional skills, if discussions between policymakers and management show this is warranted.	Develop monitoring system to ensure agreed-upon activities take place.
	Program management expects activities that policymakers leave out or de-emphasize.	M	Bring policymaker's and program management's expectations into alignment, to reflect changes in or maintenance of activities.	Maintain emphasis on activities that policymakers had not expected, if discussions between policymakers and management show this is warranted. Otherwise, drop or de-emphasize these activities.	Develop monitoring system to ensure agreed-upon activities take place.
Do policymakers and program managers agree on the immediate outcomes and on the longer-term impacts of the program?	Policymakers and program management differ in expected outcomes, which policymakers barely consider, and in desired longer-term impacts. To some extent, these reflect differences in emphasis: upon immediate use of functional skills, career development, organizational change, and so on. In some instances, expectations appear to be in direct conflict.	H	Determine where differences reflect real conflict, or merely emphasis and foresight. Based on agreed-upon changes in activities, alter expectations so they are mutually acceptable.	Add, drop, or modify the emphasis on activities linked to agreed-upon clarification in expected outcomes and impacts.	Develop monitoring system to ensure agreed-upon activities take place.
Do policymakers and program managers agree on the evidence needed to show that the program is succeeding?	Policymakers want hard evidence of impact, while program managers are more willing to accept reaction data and program audits.	M			Once expectations about activities, outcomes, and impacts are in reasonable alignment, outline measures and comparisons associated with each, and the different uses to be made of evidence by these two user groups. Select measures that impose no unacceptable cost burdens and are critical to both.

EXHIBIT 31(I)

NEP/Police Management Training

IS THE PROGRAM VALID?

Basic Question: Does the program, as it actually operates, validly represent program management's expectations?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does the actual program use the resources and serve the target population that program management intended?	Actual instructor selection criteria differ from those intended, perhaps reflecting divergent concepts of the program's general expectations.	M	Bring program manager's and operator's expectations into closer alignment about instructor selection.		Develop matrices displaying selection criteria against specific program expectations.
	The program serves a population lower in rank and responsibility than intended by management.	H	Bring program manager's and operator's expectations about target population characteristics into closer alignment, shifting where documented demand warrants.	Depending on documented demand, either tighten selection criteria while expanding outreach to attract intended audience, or gear program contents to the current lower level.	Develop procedures for estimating demand by intended and actual trainee populations, and identifying factors encouraging attendance by lower population.
	The actual program all but ignores the requirements in management-imposed curricula, viewing requirements as both out-dated and unenforceable. When management's auditors come on site, instructors echo buzz words from the requirements to show compliance.	M	Bring program managers and operators into agreement that curriculum requirements that are responsive to changing needs are worth enforcing.	Adjust program content in response to periodic review.	Develop procedures for periodic and systematic review of curricula requirements for continuing relevance.
Does the actual program represent the activities, and the relative priorities among activities, as set forth by program management?	Program management expects activities the program operators have left out or de-emphasized. This applies especially to program management's intended focus on functional skills.	M	Bring program manager's and operator's expectations into alignment, to reflect changes or maintenance of activities.	Increase emphasis on activities left out, if discussions between program managers and operators show this is warranted.	Develop monitoring system to ensure agreed-upon activities take place.
	Program operators expect activities program management left out or de-emphasized. This applies to both the program operator's emphasis on informal processes, and to the operator's broad focus on management philosophies and styles.	H	Bring program manager's and operator's expectations into alignment, to reflect changes in or maintenance of activities.	Maintain emphasis on activities that program management had not expected, if discussions between program managers and operators show this is warranted. Otherwise, drop or de-emphasize these activities.	Develop monitoring system to ensure agreed-upon activities take place.

EXHIBIT 31(2)

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	Program management either calls for traditional instructional methods or overlooks technique entirely, while program operators stress gearing technique to the desired type and level of learning.	L	Clarify how immediate outcomes and later performance impacts relate to current and alternate instructional techniques, and bring expectations into alignment.	Maintain instructional techniques that best support mutually-accepted intended outcomes.	Develop matrices for illustrating how topics, instructional techniques, and intended outcomes are linked.
Does the actual program direct its resources toward the immediate outcomes and long-term impacts intended by management?	Program managers and operators differ in expected outcomes and desired longer-term impacts. Program managers emphasize attainment of functional skills and ability to use a common language of management concepts, while program operators stress changes in attitude and orientation toward management roles.	H	Determine where differences reflect real conflict, or merely emphasis and foresight. Based on agreed-upon changes in activities, alter expectations so they are mutually acceptable.	Add, drop, or modify the emphasis on activities linked to agreed-upon clarification in expected outcomes and impacts.	Develop procedures for assessing what program emphasis the target population needs. After adjusting activities, develop a monitoring system to ensure agreed-upon activities take place.
	Program management expects trainees to leave training with a broader familiarity with prevalent management philosophies, and more capable of adjusting to the philosophy of his chief, but directed toward one or another philosophy no more than before training. Program operators aim to change management philosophies in a clear direction, to influence the management philosophies of trainees' agencies.	M	Determine if the program operator's emphasis is acceptable to management, and bring expectations into closer alignment.	Modify activities to reflect changes in expectations.	Develop a monitoring system to ensure agreed-upon activities take place.
	Management expects immediate implementation, while program operators stress incremental implementation appropriate to agency circumstances and the trainee's own sense that a particular change is valid.	M	Determine if differences stem from discrepant expectations or from differential sensitivity to external factors beyond the program's control upon which implementation depends. Bring expectations into alignment accordingly.	Modify activities so trainees recognize realistic implementation schedules.	Develop matrices showing desired types and levels of implementation, related activities, and factors beyond the program's control upon which implementation depends.
Does the actual program reflect management's views on the evidence needed to show that the program is working?	The program's ways of measuring knowledge gained in the program and its required levels of mastery are significantly lower than those desired by program management.	M	Determine if differences stem from discrepant expectations or from differential sensitivity to costs and potential uses of information. Bring operator's expectations into alignment with management's, if discussions show this is warranted.	Modify activities to produce agreed-upon levels of mastery.	Develop matrices showing alternate measures and comparisons, the intended uses of information, and the costs associated with each.

EXHIBIT 31(3)

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	The attitudinal measures adopted by the program are not in line with the specific, although partially articulated, intended changes in management style.	L			Identify instrumentation that appropriately measures the variables of interest.
	Program management wants to see hard evidence of program impact on particular aspects of job behavior, but program operators cannot agree on changes in behavior that would reasonably flow from the program. Program operators are hence content, or at least resigned, to obtain follow-up measures of trainees' postprogram reaction, or accept spontaneous testimonials to training's utility.	M	Shift program operator's implementation expectations away from course content and toward the individual trainee's documented changes in prioritized training needs, and the individual's implementation of a personalized action plan. Bring managers and operators into alignment about individualized implementation expectations.	Incorporate the development of end-of-course action plans into the program as a final training exercise.	Develop procedures for generating pre- and/or post-trainee profiles of management training needs. Develop procedures for following up on personalized end-of-course action plans.
	Program management wants to see evidence that offered training is relevant to job needs, but the program views the selection of trainees in need of available training as the responsibility of the trainee's agency.	H	Bring program managers and operators into closer alignment about the amount of control the program can exercise in selection and about the program's responsibility to determine needs and adjust the program to them.	Expand the program's role in selection, adjust activities to meet actual needs, and/or incorporate procedures for measuring changes in management training needs.	Develop procedures for generating pre- and/or post-trainee profiles of management training needs.

EXHIBIT 32(1)

NEP/Police Management Training

IS THE PROGRAM PLAUSIBLE?

Basic Question: Is it plausible that the program will accomplish its purposes?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does the program possess resources adequate to achieve its purposes?	The program does not select instructional staff on the basis of either familiarity with the substance of the course curricula or ability to use current training principles.	M	Determine if inappropriate selection criteria result from policy or tradition, and alter expectations about necessary trainer credentials accordingly.	Make trainer selection criteria job relevant.	Develop matrices showing the program's objectives and the priority of instructor selection criteria against them.
	The program has little or no control over selection of trainees who need training and screening out those who do not.	H	Alter the expectations of program staff and police agencies about the program's role in selection.	Expand the program's role in selection.	Develop procedures for measuring the pretraining needs of applicants, as a basis for screening out those who do not need training.
	The extent and duration of the intervention the program provides is insufficient to effect the type and level of desired outcomes and impacts. The program offers "too little, too late." This is especially true when the program aims at attitude change, since short-term changes are unrealistic even in most cogent programs.	M	Clarify the types and levels of desired outcomes and impacts justified by the current program, the changes in the program needed to meet expectations, and the realistic time-frames for current desired implementation.	Expand selected activities to make more plausible the accomplishment of priority expectations. Coordinate installments of a career development sequence to build incrementally toward expectations.	Develop matrices showing desired outcomes and impacts, their relative priority, the activities related to their accomplishment, and the plausibility of expectations given current activities.
Does the program effectively mobilize its resources toward its objectives?	The program retains control over trainee selection, but selects on the basis of formal factors, rather than demonstrated individual need.	H	Alter expectations of program staff and police agencies about either the program's selection criteria and/or about the program's role in documenting changes in training needs.	Select trainees based on demonstrated individual need, and/or incorporate procedures for measuring changes in management training need.	Develop procedures for measuring the pretraining needs of applicants, as a basis for either screening out those who do not need training and/or generating pre- and/or post-trainee profiles of management training need.
	The program relies on traditional lecture and discussion instructional techniques that are inadequate to its purposes of initiating attitude change or providing genuine experiences with alternate management systems.	M	Alter expectations of program staff about difficulties in departing from a lecture and discussion format.	Expand use of experiential instructional techniques.	Develop matrices showing expectations and the extent to which they might plausibly be accomplished using a range of techniques.

EXHIBIT 32(2)

Subquestion	Roadblocks to Evaluability	Fre-quency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	The program uses experiential techniques, but does not debrief trainees on the concepts behind training experiences. As a result, trainees cannot adequately explain experiences or apply relevant concepts to actual job situations.	L		Discuss the concepts behind experiential exercises after their completion.	
	The program gives instructors maximal discretion and makes few efforts to coordinate their activities. As a result, except where trainers work together on their own initiative, each trainer provides an independent intervention, often duplicating effort, failing to effectively build on other presentations, or unknowingly contradicting them. Trainers point trainees in every direction, and no direction at all.	H	Convey to instructional staff the priority to be placed on coordinating instruction.	Periodically conduct meetings of instructional staff, and monitor staff presentations to increase coordination.	Develop procedures for structured observation and assessment of the instructional process.
	Management's systems for auditing compliance of curricula with requirements do not specify the criteria for program assessment.	H	Convey to programs the criteria against which the instructional process will be audited.		Identify criteria and develop measurement systems for auditing program compliance.
Does the program have positive evidence, or at least no negative evidence, that it is achieving its intended outcomes and impacts?	End-of-course exams do not provide adequate evidence of learning, because they do not treat the outcome variables of interest, do not require a sufficiently stringent level of mastery, or both.	H	Convey to trainers and trainees that exams will be better directed toward intended outcomes and require a higher level of mastery.	Administer exams that are better directed toward the program's intended outcomes and require a higher level of mastery.	Develop procedures for constructing exams directed toward clearly defined outcomes and requiring an appropriately challenging level of mastery.
	The program has systematically collected no evidence of implementation or it has evidence that trainees find they lack the authority or leverage to implement recommended practices on the job.	H		Conduct comparable training at all management levels, to ensure the several levels of police management are on the same wavelength are hence willing to support each other.	Develop follow-up evaluation procedures to determine if the obstacles to implementation are internal to the trainee (not enough knowledge of implementation stragety) or in the external organization (resistant to concepts introduced by lower levels of management).
Lacking substantial evidence that the program is achieving its purposes, can the program document that it offers training that trainees actually need?	Programs are not designed on the basis of data showing areas where individual and organizational performance is unacceptable.	H	Establish the expectation that programs will be based on solid evidence of performance needs and not on general views of topical interests.	Develop programs based on data showing where individual and organizational performance is unacceptable.	Develop procedures for identifying individual and organizational performance deficiencies.

CONTINUED

2 OF 4

Subquestion	Roadblocks to Evaluability	Fre-quency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
	Trainees are not selected by either programs or their agencies based on demonstrated individual need.	H	Alter expectations of program staff and sending agencies about the program's interest in individualized selection.	Require agencies sending trainees to document individual needs, or select trainees based on individual needs.	Develop procedures for measuring the applicants pretraining needs for use by the program or by agencies sending trainees in the selection process.
	The program does not document the management training needs of trainees prior to or at the start of training.	M	Convey to trainers and police agencies the program's expanded role in documenting pretraining needs.	Incorporate procedures for measuring the pretraining needs of trainees.	Develop procedures for measuring the pretraining needs of trainees.
	The trainees who actually arrive for a course are often grossly dissimilar from the population for which the course was designed and for whom it would presumably be relevant.	M	Convey to sending agencies the importance of compliance with selection criteria, noting that criteria will either be tightened or shifted to the actual audience, whichever is warranted.	Depending on documented demand, either tighten selection criteria while expanding outreach to attract the intended audience, or alter selection criteria and contents to the actual audience.	Develop procedures for estimating demand by intended and actual trainee populations, and identify factors encouraging attendance by an inappropriate population.
	The program does not adjust training contents and the required level of mastery to meet the specific capabilities and needs of the immediate audience.	L	Convey to trainers their obligation to address the needs of the actual audience.	Incorporate procedures for determining the needs and capabilities of the actual audience, and adjust contents and required mastery levels accordingly.	Develop procedures for systematically determining the needs and capabilities of the actual audience, and for monitoring how trainees adjust the course to them.
	Departments expect trainees returning from the program to display skills that the program is not geared to produce.	M	To the extent it will not be politically damaging, convey to sending agencies a realistic view of what to expect from returning trainees.	Maintain the program's current focus while making trainees aware their agencies may have unrealistic expectations, or modify the program's focus to better fulfill agency expectations.	Develop procedures for involving agencies in the initial determination of training needs.

EXHIBIT 33

NEP/Police Management Training

IS THE PROGRAM'S DATA SYSTEM FEASIBLE?

Basic Question: Are the program's intended means for demonstrating success feasible?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Is the program's data system free from unacceptable cost burdens?	Consideration of approaches for providing evidence of success, other than routine reaction surveys and final exams, is unfeasible, because trainers lack skills in evaluation and the training agency lacks funds to hire outside consultants.	H	Convey the expectation that the instructional staff will include individuals with evaluation skills.	Take evaluation skills into account when selecting trainers.	Develop personnel rating scales that give prospective trainers credit for evaluation experience and capability.
	The intended uses of evaluations, as a source of program justification and guidepost for changes in program components, are far exceeded by the costs for a full evaluation of a program's payoff.	H	Convey that evaluation systems of varied costs exist, and that a system should be chosen that provides a necessary and sufficient levels of confidence in results. Reduce expectations that any evaluation less than the ideal is not worthwhile.		Identify alternate measurement systems and their costs.
Is the program's data system free from unacceptable political burdens?	Trainees resist data collection out of fear they will be judged competitively against their peers.	L		Communicate to trainees that they will not be judged competitively against their peers.	
	Unions discourage cooperation with follow-up data collection efforts, arguing that confidentiality will be violated by incorporation of data into performance appraisals.	L	Clarify the program's basis for granting confidentiality and the unions' grounds for expecting confidentiality will be violated.	Communicate to trainee that their confidentiality will be honored.	Develop clear confidentiality guidelines on the use of evaluation data.
	Program operators resist data collection for fear of providing potentially damaging evidence to funding and oversight agencies, thus jeopardizing program support.	M	Bring program managers and operators into closer alignment about the evidence needed to demonstrate success and the appropriate users of evaluation data.		Develop causal flowcharts representing program manager's and operator's views of how the program's resources and activities are linked to outcomes and impacts, and identify similarities and differences.

EXHIBIT 34

NEP/Police Management Training

IS THE PROGRAM'S DATA SYSTEM RELIABLE?

Basic Question: Does the program's intended data system have provision for repeated observations?

Subquestion	Roadblocks to Evaluability	Frequency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does the program rely on instrumentation that has established test/retest reliability?	The program administers tests in self-assessment activities or during follow-up evaluations that lack established reliability.	L			Identify tests that better meet needs for reliability
Is observation of program activities structured so independent observers may obtain similar results?	Procedures for auditing programs are not structured and do not require assessment against clear criteria, making the auditor's observations dependent on his idiosyncrasies.	M	Convey to instructional staff that program audits will involve structured observations.		Develop auditing procedures that require structured observation and assessment against clear criteria.
	Instructors assess performance in skill demonstrations or structured experiences based on personal judgment, without benefit of clear criteria and standards for structured observation.	L	Convey to trainees and instructional staff that instructors will base their assessments of trainees upon standards for structured observation while using clear and explicit evaluative criteria.	Incorporate standards for structured observation into the training process.	Develop standards for structured observation of trainees by instructors and clear evaluative criteria.

EXHIBIT 35

NEP/Police Management Training

ARE THE PROGRAM'S INTENDED USES OF DATA PLAUSIBLE?

Basic Question: Are management's intended uses of evaluation evidence under its control?

Subquestion	Roadblocks to Evaluability	Fre-quency	Adjustments in Expectations	Adjustments in Activities	Adjustments in Information Systems
Does the intended data system allow management to identify deviations from desired results to a convincing degree?	Reaction surveys, as the main sources of evaluation evidence, show little about knowledge, much less about actual performance.	H	Bring program managers and operators into closer alignment about the need for collecting evidence directly related to immediate learning.		Identify alternate sources of evidence for attainment of immediate program outcomes and their associated costs.
	Follow-up evaluations are premised on no clear concept of the types of behavior to be impacted, much less their expected performance levels.	H	Bring program managers and operators into closer alignment about the types of behavior to be impacted and desired performance levels.		Identify alternate sources of evidence for agreed-upon intended impacts and their associated costs.
Does program management possess the influence and interest to alter program activities to correct for shortcomings in desired results?	The low continuity in instructional staff and substantive contents from one time a course is offered to the next limits the chances that assessments of one session will have useful carryover implications for later sessions.	M	Reduce expectations that evaluation evidence will have substantial implications for later courses, or convey the value of increasing program continuity.	Increase continuity in instructional staff and substantive contents.	If discontinuity is unavoidable, identify evaluation measures that meaningfully apply to the individual course.
	The range of actions management takes on the basis of program evaluations is restricted to altering program components to improve trainee reaction and minimally involves efforts to upgrade performance.	H	Clarify management's expectations, to determine if its intended evaluation system really corresponds with expectations. Clarify the extent of management's power to take stronger action.		Based on clarification of management's intended outcomes and impacts, identify alternate measurement systems and their associated costs. Develop matrices showing the types and levels of evidence needed to take stronger action.

policymaker, program manager, and program operator. A policymaker has power to legislate approval for initiation of program development with a jurisdiction's or organization's support. A program manager serves in a direct oversight capacity by controlling the funding flow to a program and/or determining the program's intended activities and orientation. The program operator translates the program manager's partially defined intentions into practice and actually delivers the program that the policymaker legislates and the program manager oversees. The three groups often overlap in practice because the same person may be a critical player in two or even three of them.

B. EVALUABILITY CHECKLIST AS A SELF-ASSESSMENT TOOL

The simplicity of the Evaluability Checklist is its strength. It reduces many complex concepts into one practical self-assessment tool. It not only helps assess the current condition of a program but also points out actions that may be taken to manage the program better. The Checklist can be used in several ways, but the following tested procedure seems most effective:

- Read the headings, or overall questions, at the top of all seven exhibits.

- Go back to the first exhibit and read the subquestions. Do this for all seven exhibits.

- Go back to the first exhibit again and read all the roadblocks listed under the first subquestion. Circle all that apply in any degree to the program under consideration. Continue this procedure for each subquestion in all seven exhibits.

- Go back again to the first exhibit. Now rate how important it is to reduce the disabling effect of each circled roadblock. Use any appropriate rating system (e.g., H, M, L; 1-5).

- Go back over the roadblocks given the highest importance rating. Read the potential adjustments that correspond to these roadblocks. Circle the adjustments that appear feasible and cross out those that do not.

- Review the circled adjustments. Assess the amount of effort that they will require. If unsatisfied with the effects that these adjustments are likely to have upon the program, go back over the roadblocks given the second highest importance rating. Circle those that appear feasible and cross out those that do not. Repeat this procedure as long as necessary.

- Bring together the prioritized potential adjustments into a strategy for program improvement.

- After efforts to implement the strategy, start the assessment procedure over again. Focus on the roadblocks that were central to the strategy but do not altogether ignore other roadblocks that may have changed in importance. Repeat the entire procedure periodically.

Although the simplicity of this self-assessment process is its strength, we do not intend that it be uniformly useful. First, the Checklist could not exhaustively state key subquestions, roadblocks to evaluability, or potential adjustments. On the contrary, we hope it will suggest ways to expand on relevant parts. Second, we recognize that the roadblocks do not have uniform relevance to all programs. Some of the roadblocks, in fact, are clearly incompatible with others and, hence, could not be observed in a single program. Depending on a program's immediate circumstances, a given roadblock and its adjustments will appear platitudinous, self-evident but painfully so, insightful, or provocative. A program should selectively identify and prioritize the roadblocks that affect its operation. Third, we realize that the program adjustments we list can be enormously complex and that our abbreviated descriptions of these adjustments seem "easier said than done." There is no question, for example, that attempting to persuade a key player that his expectations for a program are out of line with needs can be a long and difficult process. Precisely because the necessary adjustments are so program specific, we have only tried to suggest here the broad scope of each potential adjustment.

C. NATIONAL ASSESSMENT OF PROGRAM EVALUABILITY

The strength of the Evaluability Checklist as a self-assessment tool, its simplicity, is probably the Checklist's weakness as a national assessment of program evaluability. Based mainly on direct on-site observations, the Checklist indicates the <u>frequency</u> with which programs confront each roadblock. The <u>importance</u> of each roadblock, in contrast, is specific to each program and is harder to judge. As a national assessment of program evaluability, the Checklist lets us reach these conclusions:

- In one or more ways, most programs are likely to be less than optimally evaluable in relation to each criterion.
- On virtually all seven criteria, nearly all programs face multiple roadblocks to evaluability.

Police management training programs are not alone in facing evaluability problems, despite the impression this report may convey. Two easily obtainable brief reports, one by Jackson and Kulp and the other by the U.S. Office of Personnel Management (OPM), show otherwise. Jackson and Kulp relate the constraints in evaluating the management training programs for which AT&T annually spends over $100,000,000. Certain of their findings parallel our own. They find that:

- Clear definition of desired performance seldom exists, making testing or evaluation difficult to impossible.
- Both the need for training and its contents are often mandated, making performance or needs analysis prior to training--and evaluation later--of limited use or interest.

- It is particularly difficult to assess the value of management training in measurable terms, due to differences between management and technical performance.6/

The OPM training demonstration project deals with several issues easily related to our own findings. For example, programs inadequately define their purposes and intended outcomes. Data systems lack reliability, because evaluation is conducted as a one-time terminal event rather than a continuous process. Expectations are implausible because trainees are selected without regard for individual need and because training is perceived as the sole solution to larger problems. Data systems are infeasible because most trainers lack evaluation skills and because little cost data exist for the evaluation of training.7/

A third and ongoing study, conducted by the American Assembly of Collegiate Schools of Business (AACSB), is also related to our study. It attempts to develop ways to assess and accredit business programs on the basis of how its graduates perform. This focus on performance measurement represents a sharp departure from precedent. Business schools have traditionally been assessed by analyzing faculty credentials, not by comparing outputs against performance standards.8/ So, in the effort to make management training programs more evaluable, we have the company of industry, government, and academia.

6/ S. Jackson and M. J. Kulp, "Designing Guidelines for Evaluating the Outcomes of Management Training," in Peterson (ed.), <u>op</u>. <u>cit</u>., pp. 1-42. Jackson and Kulp note several differences between management performance and technical performance that might influence evaluation. Management performance is more likely to be covert and harder to describe, observe, and measure than technical performance. Its standards are less likely to be defined or agreed upon. It is less likely than technical performance to be sequential and more likely to be of a problem-solving nature. Management outputs are not as clearly related to specific tasks and not as easy to judge for quality. Management performance more often requires responding to unpredictable, unscheduled demands.

7/ See the Office of Management and Budget (OMB)-commissioned <u>Report of the Training Evaluation Demonstration Project: Issues and Recommendations</u> (Washington, D.C.: U.S. Office of Personnel Management, Training Leadership Division, 1979).

8/ Nancy L. Ross, "Business School Ratings Studied," <u>Washington Post</u>, December 28, 1978, Section H, pp. 1-2. The referenced AACSB study attempts to develop ways of assessing and accrediting programs on the basis of how graduates perform. The initial controversy faced by the AACSB accreditation research project revolves around the appropriate measures of performance: career progression? technical skills? endurance and drive level? In the turmoil, some educators facetiously advanced pushups as the best single measure of drive level or endurance. Others endorse career progression as a measure of work motivation, especially those from institutions known for their graduates' rapid progression. An initial important conclusion of the study was that expected levels of competence should vary depending upon the roles that a school pursues. As the AACSB study moved into its second phase in early 1980, it began to examine the costs of

*   *   *   *

The assumption behind this chapter is that the evaluator should initially spend time in identification of roadblocks to evaluability and ways to reduce their disabling effects. This will produce information about program performance that program managers and operators can interpret meaningfully and to which they will be attentive. The next chapter looks at approaches for taking the next step, an evaluation of the program's effects on job behavior.

---

using output standards in schools, the impact of output standards on the academic environment, the feasibility of using industrial measurement techniques in academia, the ways of affecting personal characteristics in curriculum-oriented activity, the effects of requiring differential levels of output depending on the roles for which students are being prepared, and the value-added concept: the effect of an educational program on the student's managerial competency.

Six: Single Program Evaluation

Development of an appropriate evaluation design for a given program or course calls for careful consideration of the issues and application of the principles set forth in several earlier chapters and related appendices. Chapter Five, in particular, detailed the preliminary questions that one needs to ask to assess a program's evaluability. The seven questions deal with the completeness and consistency with which the program's expectations are defined, the acceptability of these expectations to policymakers, the validity of the actual program as the fulfillment of these expectations, the expectations' plausibility, the feasibility and the reliability of the program's intended data system, and the plausibility of the intended uses of evaluation data. Each question was refined into a series of pinpointed subquestions for helping individuals with evaluation responsibilities not only to determine the types of evaluation to which the program is immediately amenable but also to identify ways in which the program needs to be modified to make it more evaluable in the future.

For the purposes of developing an appropriate evaluation design, we underscore the importance of asking a sequence of questions like those in Chapter Five before designing and implementing the evaluation. Other useful references can also be consulted in determining whether and how to evaluate, including A Process for the Evaluation of Training, developed by the U.S. Office of Personnel Management (OPM, formerly the Civil Service Commission).9/ Based on the results of this preliminary assessment, original evaluation concepts may be clarified, modified, or radically altered so as to make and monitor needed program adjustments. Once this process has brought the program to a condition of acceptable evaluability, Appendix 6 can help to identify precisely how the program's results ought to be measured. Exhibit 36 lists the range of questions that a program evaluation might involve.

This chapter's purpose is to introduce three evaluation approaches that have wide applicability and require relatively low investment. These approaches involve a follow-up on implementation of personal end-of-program action plans, an examination of training's effects on proficiency, and a simplified cost-benefit analysis. All three approaches can be used to answer the question: Does training make any difference in later job behavior? Each deals with this question in a different way. The action plan follow-up approach asks: Do trainees successfully carry out on their jobs the intentions they carry away

---

9/ A Process for the Evaluation of Training (Washington, D.C.: U.S. Civil Service Commission, Training Leadership Division, April 1978). A copy may be obtained from the U.S. Office of Personnel Management, Workforce Effectiveness and Development Group, Office of Consulting Services, Training Consulting Division, 1900 E Street, N.W., Washington, D.C. 20415

EXHIBIT 36

NEP/Police Management Training

RANGE OF QUESTIONS FOR PROGRAM EVALUATION

Do trainees leave the program genuinely satisfied that they have learned something applicable to their jobs?

Do trainees return to their jobs with more knowledge than they had before training in areas on which the program focused, such as minimum required duties and responsibilities, management principles and practices of a particular type, problems in and preconditions for implementation, and so forth?

By the end of the program, do trainees have greater self-knowledge of their personal management orientations and styles?

By the end of the program, have trainees begun to internalize their experiences with alternate management systems?

Before they return to their jobs, are trainees able to analyze hypothetical management problems in terms of alternate management systems?

Do trainees leave the program motivated to implement what they have learned and with reasonably clear plans showing what they expect to transfer, where they expect to gain support, and what obstacles they are likely to encounter?

Are the practices that trainees say that they want to try on their jobs at the end of the training in line with those that the program intended trainees to implement?

Do graduates return to their jobs with greater confidence and boosted morale?

Back on the job, do graduates retain program concepts and use the common language conveyed in training to analyze department problems?

Do graduates perform their duties in increased compliance with departmental regulations and policies?

Do graduates make genuine efforts to try on their jobs some of the principles and practices learned in training? To what extent are these efforts successful?

Can graduates identify situations that are appropriate for the implementation of new behaviors?

When graduates meet with success in trying new behaviors, do they gradually internalize their experiences with alternate management systems, and do their attitudes change as a result?

Do graduates display greater proficiency after completion of training in the areas where they were deficient before training?

Do graduates increase their participation in career development activities, including advanced education and training?

Do graduates experience more accelerated career development than before training?

Do graduates use the "network" of graduates to solve management problems encountered on the job? Does use of the network change their decision-making styles?

Does the high morale that graduates carry back to their jobs spread to officers within their units and have a positive effect on unit output?

Are the benefits to the department of the graduates' greater proficiency worth the costs of training?

Does the presence in a department of a high percentage of similarly well-trained managers affect the department in terms of output, smoothness with which large-scale organizational change takes place, agency reputation, and so forth?

from training? The proficiency analysis asks: Does training improve the trainee's proficiency in ways that the organization can use? Simplified cost-benefit analysis asks: Is training a reasonable investment to make, based on its return to the organization in increased job proficiency?

The three approaches outlined here do not, by any means, exhaust the ways to examine training's effects on job behavior. A really rigorous evaluation could involve direct structured observation of work samples. At the opposite extreme, evaluation could simply test for retention of course learning several months after return to the job. The three described approaches are highlighted for two reasons. First, they exemplify the most advanced state of the art in evaluation of management training. All three have shown their usefulness in industry and government, but none has been extensively used to evaluate police management training. Second, supplementary descriptions of the three, if they are needed in translating general descriptions of the approaches into practical procedures, are readily obtained.

The following discussion draws heavily on two sources of information: works published by OPM and our experiences during this study in experimenting with and testing the usefulness to law enforcement of the first two approaches. We rely on the OPM materials for two reasons. First, these provide excellent guidelines, complete with all needed directions, sample scripts, forms, and instruments. Second, programs operated by state or local governments in most OPM regions may obtain assistance in implementing these approaches from the regional office at nominal cost. We balanced the OPM perspective against our own experiences because we wanted to ensure that these approaches were capable of generating information useful to police management training.

A. ACTION PLAN FOLLOW-UP APPROACH

The Action Plan Follow-Up Approach (APFUA)--which OPM calls the Participant Action Plan Approach (PAPA)--addresses the central question: Do trainees successfully carry out on their jobs the intentions they take away from training? It may be used to examine a score of more pointed questions, listed in Exhibit 37, such as:

- To what extent do end-of-course intentions lead to implementation of new behaviors?
- Does the course result in new behaviors other than those expressed in end-of-course intentions?
- What portion of new behaviors reflects course learning objectives?
- In what ways does changed job behavior impact on trainee organizations?
- To what extent have trainees encountered problems in implementing new behaviors?
- What aspects of the course produce the greatest changes in behavior?

EXHIBIT 37

NEP/Police Management Training

SAMPLE QUESTIONS THE ACTION PLAN APPROACH CAN ADDRESS

What types of intentions did trainees carry away from training?

To what extent did trainees formulate end-of-training intentions?

To what extent did trainees leave the course with intentions that reflected course objectives?

To what extent did these plans lead to implementation of new behaviors?

To what types of new behaviors did they lead?

To what extent did the course result in new behaviors that were unplanned, in that they were not expressed in end-of-course intentions?

What types of unplanned new behaviors were implemented?

To what extent were changes in trainee behavior planned?

To what extent did new behaviors, whether or not expressed in end-of-course intentions, reflect course objectives?

What types of impacts did these new behaviors have on trainees' agencies?

Did trainees think that these new behaviors had a positive, negative, or mixed impact on their agencies?

To what extent did trainees encounter problems in trying these new behaviors?

What types of problems did they face?

Which aspects or parts of the course--defined by course objectives, modules, or substantive groupings--were most prominent in end-of-course intentions?

Which aspects or parts were reflected in changed trainee job behavior?

Which aspects or parts yielded the greatest number of positive impacts on the trainees' agencies?

To what types of outcomes did different aspects or parts of the course lead?

Which aspects or parts tended to produce problems in implementing new behaviors?

To what types of problems did different aspects or parts of the course lead?

To what types of non-behavioral changes and planned behaviors did the course lead?

How high were the reported levels of new knowledge and attitudes and of planned but untried behaviors relative to the number of new behaviors?

What types of changes did trainees recommend for the overall course?

APFUA is a flexible approach that has been refined into a fine science but really boils down to two tasks. The first is to guide trainees through the development of personal end-of-course action plans. The second is to follow up after a few months to see how trainees have fared with their plans, identify other new behaviors implemented as a result of training, determine the impact of new behaviors on the organization, identify implementation problems, and so forth. Either one of these two tasks--the action plan development and the follow-up on implementation--could be done independently but without the effects of the two combined.

This approach has two distinguishable functions. The first is to motivate trainees to transfer course learning to their jobs, i.e., to change trainee behavior. The second is to provide a framework for evaluating the course's influence over new behaviors. Both uses are documented in OPM's and our experiences with the approach. Writing action plans gets trainees more involved in the course, stimulates them to think about transfer of course contents, and makes them be more specific about how to apply course contents to their actual job situations. The action plan itself implies a commitment; expectation of a follow-up intensifies this implied commitment by reminding trainees to work on their plans and by making them feel more accountable for at least trying implementation. Using action plans as the baseline for the follow-up provides an explicit starting point for discussion of behavioral change and a point of comparison for assessing the degree of change.

Although a fair number of programs conduct action plan exercises and appreciate the action plan's role in instigating change, few recognize its evaluation role and try to follow up on implementation. From our experiences, APFUA evaluation data may be used effectively in several ways, such as to document course outcomes for marketing purposes, to identify ways in which course contents or selection procedures require change, to provide narrative support to quantitative evaluation data, and to establish the need for and focus of full-scale evaluation.

There are certain preconditions and requirements for use of APFUA that should be borne in mind. The OPM authors stress that an approach like this is especially suited to evaluate programs that lack either pretraining measures of knowledge and skill or easily measured learning objectives. At the same time, it should not be used when other approaches can answer questions about behavior more directly or as the only data source for determining if a course should be discontinued or radically altered. The APFUA approach is generally used to collect information from the perspective of the program manager or operator, but it may also be followed by sending agencies to determine if training offered is appropriate to their needs. Because the approach allows for an implementation period, it requires, like any follow-up, waiting long enough for new behaviors to appear. It requires certain staff skills as well. The individual who introduces the action plan concept and guides the action plan exercise must have the skills to explain how to develop an action plan thoroughly and to motivate trainee involvement in plan development. The individual who collects follow-up data must have excellent interviewing skills to be able to probe for specific instances of behaviors, impacts, and problems and to judge whether trainees are being factual. The individual who analyzes the data must possess basic computational and analytical skills for working primarily with qualitative data. The

course structure and orientation need not be significantly changed to incorporate APFUA, although they may be. Minimal changes need to be made if APFUA's primary purpose is evaluative. Near the start of the course, the concept of action planning must be introduced; near the end, the action plan exercise must be conducted. More extensive changes are probably undesirable if the existing course is to be accurately assessed. To the extent that APFUA's purpose is viewed as motivational, the course's structure and orientation can and should be changed in other ways. Instructors may focus trainees heavily on application of learning, time periods may be set aside for jotting down action items, action idea development may be tied directly to course objectives or otherwise made a pervasive course activity, and so forth.

This APFUA overview contains eight major steps, starting with the resolution of preliminary evaluation questions and ending with development of a program evaluation report. Each major step has several substeps. With this level of detail, the description should generally be adequate to permit useful adaptation to particular needs. For supplementary information, the following resources will be helpful:

- OPM's handbook on the Participant Action Plan Approach[10]
- An OPM videotape describing how to use the approach
- Related articles in Peterson's anthology on management training evaluation[11]
- Sample reports from our feasibility tests with APFUA[12]

10/ Assessing Changes in Job Behavior Due to Training: A Guide to the Participant Action Plan Approach. (Washington, D.C.: U.S. Government Printing Office, 1980). Requests should refer to stock number 006-000-01130-6.

11/ R.O. Peterson (ed.), Determining the Payoff of Management Training (Madison, Wisconsin: American Society for Training and Development, 1979). See R.D. Salinger, "Measuring Behavioral Change Which Results from Training," pp 113-150, and J. Pittam, "Evaluating On-Job Performance of Participants in General Management Seminars, pp. 151-170.

12/ Sample reports may be borrowed from the National Criminal Justice Reference Service. The two available reports are entitled: "Evaluation Report: Police Executive Development Institute (POLEX)" and "Evaluation Report: Line Supervision Course." The first was conducted for the Pennsylvania State University and the second for the Southeast Florida Institute of Criminal Justice (SEFICJ) at Miami-Dade Community College.

- Opinions of the training program directors who hosted our feasibility tests[13]
- Assistance from OPM's Training Consulting Division (in Washington, D.C.) or from an OPM regional office

1. Resolve Preliminary Evaluation Questions. Before proceeding with the classroom exercises related to action plan development, there are roughly a dozen preliminary issues that first need resolution.

a. Determine what questions the evaluation will address. Out of the questions that APFUA can be structured to answer, listed in Exhibit 37, for which ones do the program's decision-makers want useful information? It is important to decide early which questions will and will not be answered by the evaluation. In reaching this decision, it is better *not* to show decision-makers the full list of questions initially because many will indiscriminately ask for the entire menu. It is better to start with a general discussion of informational needs, derive major issues from this discussion, and then verify that other potential issues are not high priority. There is little point in collecting and analyzing information that will not be used by program decision-makers.

b. Assess the relationship between APFUA and other evaluative approaches. How will APFUA tie into existing evaluation approaches? Are there more direct means for obtaining the information APFUA will generate? These questions point in opposite directions. The first expresses concern that APFUA complement the evaluation approach that the program currently uses. The second tries to establish whether other approaches can cost-effectively produce more concrete data and should be used instead of APFUA. The point of both questions is to ensure that the selected evaluation strategy has optimal prospects of generating interpretable results.

c. Assign responsibilities for APFUA tasks. Who will have responsibility for key components of the approach--specifically, the initial classroom presentation of the action plan concept, conduct of the action plan exercise, scheduling and completion of follow-up data collection, and analysis of resultant data? In some evaluations, one individual will take responsibility for all APFUA tasks, but, in most cases, tasks will be parceled out among several individuals. Typically, the course coordinator or a lead instructor introduces the concept and guides the action plan exercise. Someone who is familiar enough with the course to talk knowledgeably with trainees but who is not seen as having a stake in evaluation results conducts the follow-up. Once data have been coded,

---

13/ The POLEX and SEFICJ directors have agreed to relate their experiences with APFUA to those who are interested. The POLEX director may be reached at: POLEX, The Pennsylvania State University, Henderson Human Development Building, University Park, Pennsylvania, 16802. The SEFICJ director may be reached at: Miami-Dade Community College, Southeast Florida Institute of Criminal Justice, 11380 N.W. 27th Avenue, Miami, Florida 33167.

someone with basic course familiarity and analytic skills analyzes the data. It is useful to assign these responsibilities early, so that all involved can have a chance to contribute to formulating specific details of the process.

d. Determine how evaluators will obtain thorough familiarity with the course. How will evaluators obtain thorough knowledge of course contents? This question is important if outside evaluators collect and analyze the follow-up data. It should be decided early how they will become familiar with the course: by direct observation, review of handouts, interviews with instructors and the course coordinator, or through other means.

e. Identify course modules that reinforce action planning. Are there any course modules that could be used to drive home the importance of action planning? If there are any such modules--say, in MBO or Objective Setting--they should be identified and their instructors brought into the planning process.

f. Decide whether all instructors should be informed about the APFUA evaluation. Should all instructors be notified that the course is under evaluation and that trainees will be continually involved in assessing the transferability of course contents? The advantages of telling instructors are that they will concentrate more on how to apply course learning and will be more inclined to remind trainees periodically about the need to think about action items. The disadvantage is that the instructors' awareness produces changes in the course even before evaluation results are in. If an affirmative decision is reached, it is important to decide who will take the responsibility to inform instructors. In the event of a negative decision, it is important to assign a credible identity to any outside evaluators who may be sitting in on classes.

g. Establish the means for follow-up data collection. Should follow-up data be collected by mailed questionnaire, telephone interview, or face-to-face interview? Each has its advantages and disadvantages. Mailed questionnaires are the least costly but tend to have low response rates and do not allow detailed probing. Face-to-face interviews produce the most detailed information but are generally not feasible due to their cost. Telephone interviews can be conducted at moderate cost, allow probing for specifics, produce higher response rates than mailed surveys, and can be done at mutual convenience. We recommend telephone interviews and assume, in later discussions, that this means will be used.

h. Determine whether the full class or a representative sample will be followed up. Should the entire class be promised a follow-up contact or told that class members will be randomly selected for follow-up? The two considerations here are the class's size and the level of resources available for follow-up calls. Any class with over 20-25 members should be considered for selective follow-ups. Alternatively, the entire class might be followed up the first time a course is evaluated but only randomly selected trainees followed up when the course is again offered and evaluated. Trainees should not be promised a follow-up call if this is not really expected because broken promises discredit the entire APFUA process. If it is unclear whether resources will be available for full follow-up, then selective follow-up should be chosen with the option of reversion to full follow-up.

i. Decide whether trainee perceptions will be corroborated by another viewpoint. Should follow-up data collected from trainees be corroborated by interviews with superiors, peers, or subordinates? A second perspective on posttraining trainee behavior can be helpful, especially when the course's transferability is in doubt. There are problems with collecting data from other sources, however. First, trainees may become reluctant to cooperate if they know others will be asked to scrutinize their behavior. They might feel threatened that the follow-up will be incorporated into a performance appraisal. Second, others may not be in a position to observe trainee behaviors. They may even be the objects of certain actions or the obstacle to completion of others. These corroborative views, therefore, may be of questionable validity. We do not recommend collection of data from sources other than trainees until the process of trainee data collection has been mastered.

j. Establish the appropriate time period for follow-up contact. How soon after training should follow-up contacts be made? The follow-ups should not be made too soon, or trainees will not have had sufficient time to complete their initial implementation efforts; nor should they be put off too long because trainees tend to lose interest and because factors other than training can intervene to change behaviors and confound the evaluation results. The appropriate time for follow-ups will depend on the nature of the course, the thrust of action plans, and the immediacy of need for feedback. Follow-ups should generally be started anywhere from one to four months after training, with two months the optimal follow-up period.

k. Determine whether course objectives are clear enough to analyze data in terms of objectives. Are course objectives sufficiently clear to permit an analysis of new behaviors and other data in terms of objectives? Would such an analysis be useful? Or would analysis in terms of course modules or another classification scheme be more feasible and equally useful? There are several ways in which APFUA data may be analyzed to determine the relative influence of a course's many aspects. In many courses, objectives may be so unclear that it would be fruitless to use objectives as the basis for analysis. If a course's decision-makers are interested in analysis by objectives, it is essential to reach early agreement on the course's objectives. Alternatively, a decision may be reached to analyze new behaviors in terms of course modules or natural behavioral groupings.

l. Decide whether APFUA data analysis will be primarily quantitative or qualitative. Should data analyses be quantitative, qualitative, or both? This depends on what types of information the course's decision-makers want and can use. If they are primarily interested in the extent of change, then most of the analyses will be quantitative. If they want to know the types of change, then the analyses will be qualitative. In most cases, it will be useful to provide qualitative data (on new behaviors, organizational impacts, problems, and so forth) and to compare and summarize qualitative data numerically.

m. Identify guarantees that will have to be given. Will certain guarantees be required for APFUA to be accepted? Especially when the evaluation involves more than one organization, guarantees may be required. These can

resolve such questions as: Who will assume administrative costs? Is individual, departmental, and program confidentiality assured? In what ways and by whom may the evaluation results be used? All guarantees should be placed in writing, in case key personnel happen to change jobs.

2. Introduce Trainees To The Action Plan Concept. Early in the course and throughout it, pre-action plan exercises prepare trainees for eventual development of a personal action plan.

a. Place the evaluation process in context. Early in the course, a lead instructor or the course coordinator should explain to trainees that they will later develop an action plan, the purposes of which are to help motivate them and to provide a basis for course evaluation. It should be clarified that trainee success in carrying out their action plans will be established through follow-up contact, to occur at a specified time after training. This APFUA introduction may occur during the course's orientation session or in the first substantive module, depending on personnel available for the introduction and trainee rapport. A full description of the follow-up process may be given now or deferred until the action plan exercise. If deferred, it may still be useful to note now that the development of action items will become easier through practice and after presentation of certain related modules (e.g., Objective Setting). Either APFUA's motivational or evaluative purposes may be stressed, or both may be given equal weight, as appropriate.

b. Explain the value of trainee participation in APFUA. The potential value of APFUA to trainees, to the evaluated course, and to future trainees who take the course should be stressed. This is especially important if trainee participation is strictly voluntary.

c. Provide trainees with instruction in how to develop action items. To help trainees think about application of course learning and formulate action items throughout the course, they should be instructed in how to develop action items. Broadly stated, an action item expresses a new behavior that a trainee would like to try on the job. It should include an action verb, describe a behavior that is observable (to the trainee and/or someone else), and be meaningfully related to the trainee's job setting. For a pre-action plan exercise on action items, the OPM handbook includes a set of "Guidelines for Writing Action Items" (adapted in Appendix 7) that may be distributed to trainees. This handout or something similar should be distributed, but it will not help trainees much simply to read it in class or as homework. The instructor should have trainees read it individually and then lead the development of several practice action items on the chalk board.

d. Provide trainees with action item "scratch sheets". After the lesson on how to develop action items, trainees should be given action item "scratch sheets." The top of each sheet should read, "Ideas for Action Items," and, below that, "Ideas I would like to try when I return to my job, based on what I learned during this course." At the bottom of each sheet should be a note that "You can use the course objectives, what you learn in class, the course handouts, conversations with others, etc., to come up with ideas." The purposes of the scratch sheets are to remind trainees about the action plan they

will later develop, stimulate them to jot down preliminary ideas throughout the course, and provide a single location for keeping track of action items. Each trainee should at first be given two or three sheets. Trainees should be told that jotting down preliminary ideas as they occur is just a way of not forgetting them but that, later, they will have the chance to formalize the action items of their choice in the action plan. If trainees keep handouts in a loose-leaf binder, the scratch sheets should be hole-punched and trainees instructed to keep them in the binder. Additional scratch sheets may be made available to those who fill up their original sheets by keeping a supply on a side table with other handouts or upon direct request to an instructor or course coordinator. It is important for trainees to understand that the scratch sheets will not be collected.

e. Remind trainees about jotting down action items periodically. It is important that trainees be stimulated to develop action items throughout the course so that action plans do not simply reflect the last topic they heard about before designing the action plan. At a miminum, trainees should be encouraged to think about the transferability of all course contents. Beyond that, other steps may be taken to increase trainee attention to action item development. They can be asked to contemplate new ideas outside of class and to take time at night to jot down ideas they may want to try; or brief time periods may be set aside, five minutes or so in duration, at the end of each module or half day, to give trainees a chance to ponder potential action items. In extremely long courses, it may even be worthwhile to lengthen and formalize these brief sessions into an abbreviated action plan exercise.

f. Review action items informally and give trainees individual feedback. The individual with primary responsibility for selling trainees on the action plan concept should generally not review action items formally until they have become part of an action plan. However, it helps trainees to obtain informal feedback individually on their action item ideas. This can be handled discreetly by watching trainees as they jot down items and then unobtrusively giving them feedback right away or outside the classroom later.

3. Conduct The Action Plan Exercise. The action plan exercise should produce plans that are specific enough for trainees to act on and for the evaluator to interpret and follow up. It should leave trainees assured that their confidentiality will be protected, clear on the mechanics of the follow-up, and more capable of general action planning.

a. Start the action plan exercise in one of the course's final sessions. The action plan exercise--the formalization of action items into a final plan--should start in one of the course's final sessions but not at the very end of the course. Time should be allowed after the development of action plans for review, feedback, and putting on of finishing touches. In a two-week course, for example, it would generally be appropriate to conduct the exercise near the end of the next-to-last day.

b. Reaffirm the evaluation's context. It is useful, before distribution of action plan forms, to remind trainees that APFUA has dual purposes and that any practical idea may be included in the plans, even if not drawn directly

from class lectures and exercises. It may be noted that the best ideas often emerge from discussions among officers from different departments.

c. Distribute action plan forms and explain how to complete identifying information. Trainees should initially be given two action plan forms, with additional forms available on request. Each form should contain three sheets separated by carbons. The top quarter of the sheet should contain space for basic identifying information, e.g., course title, date, name, address, and telephone number. The bottom three-quarters of the sheet, separated by a dotted line, should lead with the phrase "Ideas I would like to try when I return to my job." The right-hand side of the lower portion should have three columns for checking off the projected time frames for plan implementation, e.g., as called for, within two months, after two months. Trainees should be asked to give the address and telephone number at which they wish to be contacted during the follow-up, not necessarily a work number. (It may be pointed out that many trainees feel more comfortable when the follow-up interview is done outside work.)

d. Explain why action plan forms contain multiple copies. Some trainees are disconcerted when they find that the action plan form contains three copies separated by carbons (even more, if corroborative reports are to be solicited from other sources). While they complete the identifying information, the use of each copy should be explained. One the trainee takes along and two the evaluator keeps; of these two, one is sent to the trainee along with a reminder letter shortly before the follow-up contact and the other is kept by the evaluator to use in the interview. Because there are three copies, trainees should be asked to make sure the carbons are going through legibly.

e. Assure trainees that their action plans will be kept confidential. Trainees should be told that identifying information is required for the follow-up contact and will be used only for the follow-up; also, that collected information will be kept confidential and used only for the evaluation. It may be noted that, after the follow-up, all identifying information will be destroyed by cutting the action forms along the dotted line and discarding the upper portion. Confidentiality should not be overplayed, however, or anxiety levels will be artificially heightened. If warranted by the trainees' confidentiality concerns, a coding system may be introduced. If trainees are not concerned about confidentiality, a coding system creates the impression of "much ado about nothing." It will generally be sufficient simply to note that no information will be used for another purpose including, if asked, individual performance appraisals.

f. Explain how to complete the action plan form. The explanation of how to complete the action plan form should build on the earlier discussion of action item development. Four points about action items must be made: they should specify actual behaviors, indicate a sequence of behavioral steps that will have to be taken, realistically account for needed supports and likely obstacles, and note the time required to start the action. The differences between specific and non-specific action items may be illustrated. It may be noted that, for some complex action items, some steps may be started within two months but others are unlikely to occur until later. A second OPM handout,

"Questions About Your Action Items," can be used to discuss organizational supports and constraints that affect implementation (see Appendix 7). If this handout is used, point out to trainees that they need not answer each question directly in their plans; rather, they should consider the questions, answer them silently, and then write out action plans that build on the answers through their specificity and realism. The action plan exercise may also be centered around course learning objectives, if objectives are sufficiently clear and if doing so interests the program's decision-makers.

g. Allow sufficient time for completion of action plans. Trainees should be given about thirty minutes--more or less, depending on course length and trainee response--to complete their action plans. It may be useful to suggest that each trainee list a minimum number of items, say two or three, to ensure that no individual takes an easy out (by completing just one) and to establish some comparability among all action plans. It is also a good idea to ask trainees after a few minutes of the exercise to check their carbon copies again.

h. Provide a mechanism for review of action plans. Each action plan should be reviewed in terms of specificity, sequencing of behaviors, realism, indication of time until initiation, and general completeness. One or both of two review approaches may be used: evaluator/instructor review and peer review. Evaluator/instructor review should occur outside the classroom to make effective use of time. Peer review, which involves the random pairing of trainees and an exchange of action plans, should occur inside the classroom but should be allotted only about ten minutes. In some courses, it may be totally inappropriate to use peer review because trainees object to it. (Such objections often arise from intra- and interdepartmental rivalries that trainees carry into the classroom.) When both forms of review are used, peer review should generally occur second. Before the peer review, general group feedback may be given on the basis of the evaluator/instructor review.

i. Provide feedback to trainees about their action plans. If evaluator/instructor review is used, it is generally appropriate to employ group feedback. This involves explaining to the entire class, or to groups within the class, the types of modifications they may want to consider in finalizing their action plans. At a minimum, the coordinating instructor should analyze examples of strong and weak hypothetical action plans in terms of specificity. This may be expanded into an exercise on revising hypothetical action plans in terms of specificity, sequencing of behaviors, realism, indication of time until initiation, and general completeness. If peer review is chosen, feedback should be direct, one-on-one.

j. Allow trainees to put finishing touches on their action plans. All action plans should be returned to their authors and time allowed for application of finishing touches based on review comments and later ideas. The amount of time needed to finish action plans depends mainly on the amount and nature of feedback. Ten minutes should be sufficient for a touch-up, but more time may be required for a total rewrite. Timing of this rewrite exercise depends on the types of review used. It can come right after peer review, but more time generally must elapse between evaluator/instructor review and the rewrite.

k. Review mechanics of follow-up contact. After finalization of action plans, the coordinating instructor should remind trainees when the follow-up contact, centered around the action plan, will occur. Specific mechanics of the follow-up--reminder letters, preliminary calls to schedule an interview, and so forth--should also be detailed.

l. Give each trainee a copy of the action plan. Each trainee should be given the first carbon copy of his or her action plan. If the carbon failed to go through, the original should be photocopied and a copy given to the trainee. If necessary, explain again why the action plan contains three copies.

4. Prepare For Follow-Up Data Collection. Preparation for the follow-up contacts involves reviewing of action plans, developing questions and record sheets, mailing reminder letters, scheduling the interviews, and motivating trainees to cooperate.

a. Draft and mail reminder letters. A reminder letter should be sent to all trainees under the signature of the coordinating instructor. The letter's purpose is to alert trainees to an imminent scheduling call and to motivate them to cooperate. It should stress the value of cooperation in the follow-up to trainees, the program, and future trainees. It should specify the day or days on which the evaluators will try to reach trainees to schedule the actual interviews. The letter should be sent a week to ten days before scheduling efforts will begin.

b. Develop question lists and record sheets. Because of the strong possibility that a few trainees will insist on completing the interview during the scheduling call rather than waiting a week or so, the evaluator should prepare question lists and record sheets before the start of scheduling efforts. The interview question list should correspond with the specific evaluation questions. Let us assume we are interested in the following topics: implementation of action plans, organizational impacts, problems encountered, other new behaviors not related to action plans, nonbehaviorial changes, and course reactions. Our basic question and probe list, to be drawn upon selectively for each action item, should then look something like this:

1. New Behaviors--Have you been able to do this yet? Specifically, what have you been able to do? What steps did you go through? What did you do first to enable you to do that? What did you do next? And next? How did you carry that out? Who else was involved? Could you give me specific examples of that, specific instances when you did that? Have you done that more than once? Have you done anything similar? Is this really different from what occurred before? How?

2. Problems--Did you have any problems in carrying it out? Specifically, what were they? Can you give me an example? If you did this more than once, were the results different from one time to the next? Will you continue to do this? Do you foresee any problem?

3. Outcomes--What were the results of doing this? Specifically, what resulted? Can you give me an example? If you did something more than once, were the results different from one time to the next? Do you think the outcomes were generally positive, negative, mixed (uncertain), or would you say there was really no change? Do you think the results will continue along the same lines?

The question and probe list should also include other questions that are not directly related to action item-related behaviors. These three basic questions should be included:

4. Other New Behaviors--Is there anything else that you are doing differently on the job since attending the course that you think is due, directly or indirectly, to being in the course? (If "yes", ask questions 1 through 3.) Is there anything else not related to your action plan that you are doing differently now and that you attribute to the course?

5. Nonbehavioral Change--Are there any other ways in which you have changed that you think are due to the course but that have not resulted in doing things differently, have not led to action? Are there things you plan to do but have yet to do? Are there new things you know, or things that you feel differently about, without this affecting how you perform your job?

6. Course Reactions--Do you think the course should be changed in any way, to make it more relevant to your job needs and to what other officers still need to know? Precisely how should it be changed? Is there anything else?

Two record sheets (or data logs) should be designed for transcribing trainee responses, one for behavioral and the other for nonbehavioral data. The behavioral record sheet should include space at the top for identifying information (trainee code, department, course, date of interview, interviewer, whether or not trainee had action plan). Below that should be space for writing out the action item to which the behavioral data relate. (For information given in response to question 4, there will be no action item.) The rest of the record sheet should have space for recording four types of information: behavior (with examples), problems/concerns, outcomes/results, and judgment of outcomes. The nonbehavioral record sheet should have space for recording three types of information: knowledge/attitude changes, planned behaviors, and ways the course should be changed. Only one nonbehavioral record sheet should be needed for each trainee. Several behavioral record sheets should be printed for each trainee, however, because a separate one is completed for each reported new behavior. Potentially, several forms will have to be completed in relation to the same action item and additional forms for new behaviors that do not even relate to an action item.

c. Review all action plans. Again, because of the prospect that some trainees will insist on jumping the gun by completing the interview during the scheduling call, it is advisable that the individual assigned to schedule interviews review all action plans first.

d. Schedule time for follow-up interview. Trainees should be contacted on the day or days indicated in the reminder letter to schedule the follow-up interviews. Because of the purposes of this preliminary contact, it is preferable that scheduling be done by telephone; if this is infeasible (e.g., because trainees work nights or have unpredictable duty schedules), it may be done by letter. This preliminary contact has three distinct purposes. The first is to establish direct contact with the trainee and set a time for the interview. The second is to get a sense for the trainee's individual responsibilities, establish the likelihood of later contact difficulties, and identify back-up times for the interview. The third is to explain to the trainee the questions that the interview will cover and, thus, to show trainees that the evaluators are interested in both positive and negative reports and to stimulate trainee review of their action plans. Information related to each scheduled interview should be recorded on a trainee scheduling card. A master schedule should show all scheduled interviews.

It is desirable but not essential to avoid completing the actual interview during the scheduling call. As a rule, trainees will not be prepared to talk at the time of the scheduling call. It is generally best to defer an interview that the trainee "wants to get over with" until the time and climate needed for probing are present. If the trainee asks to complete the interview during the scheduling call and really seems prepared to do so--is unrushed, familiar with the action plan, and comfortable with any co-workers who may overhear the conversation--then it may be done now.

e. Mail verification of scheduled interviews. If time permits, each trainee should be sent a post card indicating the time scheduled for the follow-up interview.

5. Conduct Follow-Up Interviews. Conducting the follow-up interviews requires adherence to a schedule, attention to the individual trainee's action plan, continuous efforts to probe for specifics and to separate related behaviors, and prompt transcription of interview notes.

a. Review the trainee's action plan before the interview. Immediately before the follow-up interview, the evaluator should carefully review the trainee's action plan.

b. Make follow-up calls at the scheduled time. The evaluator should try to reach each trainee at the agreed-upon interview time. However, if trainees tend to work in a response mode and do not have much control over their daily activity schedules, it is likely that a percentage of scheduled interviews will not take place. In such cases, the evaluator should display continued interest by leaving a message, by giving the option of return collect calls, and by trying to find out the trainee's work schedule and best times for making contact over the next few days. Any problems and other scheduling information should be recorded on the trainee scheduling card. Rescheduled interviews should be placed on the master schedule.

c. Follow standard probes while striving to separate behaviors. The actual follow-up interview should follow the standard questions and probes outlined above or a similar pattern. The evaluator should probe for detailed

descriptions of actual behaviors and should make sure that answers are specific enough that the described behaviors can be envisioned. When a behavior sounds so complex that it probably constitutes a sequence of behaviors, the evaluator should help the trainee sort these into a sequence and then ask the basic questions--about behaviors, problems, and impacts--for each behavior in the sequence. At the same time, the evaluator should be careful to avoid giving the impression that there are any "right" answers and should not pressure the trainee to fabricate information. As the evaluator finishes covering each set of questions (i.e., 1 through 6) in relation to a given behavior, he or she should check off the question set as completed.

d. Record narrative data on record sheets. Interview data should be recorded on the behavioral and nonbehavioral record sheets, as shown in Exhibits 38 and 39, during or soon after an interview. The more complex the data, the greater the probable need to simplify and transcribe the data onto a second set of record sheets. The evaluator should not alter the data but should be sure to put each distinguishable behavior on a separate behavioral record sheet. Each specific instance of behavioral change can be identified by looking for an action verb, a time of occurrence, persons involved, and results. This can be done easily and legitimately only if the evaluator succeeded earlier in helping trainees to separate behaviors into sequences.

e. Attach record sheets to action plans and delete identifying information. The transcribed record sheets for each trainee should be attached to the trainee's action plan, thus forming a complete data set. The only identifying information that should remain is a trainee code, which should be placed on each record sheet and action plan form in the set.

6. Organize Data In Preparation For Analysis. Once data for all trainees have been transcribed onto record sheets and identifying codes have been detached, a few steps still need to be taken to prepare data for analyses. The first two steps are optional and are really ways of compensating for shortcomings in the earlier interviewing and transcription stages. As corrective measures, they should only be taken as a last resort and still cannot transform data into the ideal form for analyses.

a. Reduce transcribed data onto cards. Sometimes the quantity of data on the record sheets becomes unwieldy for coding purposes. This situation usually would not arise when the evaluator paid attention to the original transcription process and really tried to reduce data to their essence. If this was not done earlier, it may have to be done now. The evaluator should first develop a list of variables to be coded and create a coding format. Each behaviorial record sheet should then be condensed onto a single index card. Three-by-five-inch and four-by-six-inch index cards work nicely for this purpose.

There are clear advantages to reducing data onto cards. There will be less information to rewrite repeatedly later in analyzing trainee responses. Index cards may also be shuffled and sorted somewhat more easily than record sheets when doing certain analyses. This entire step should be avoided if possible, however, because it duplicates effort. In addition, if data are condensed long after they were originally recorded, they may inadvertently be distorted.

EXHIBIT 38

NEP/Police Management Training

BEHAVIORAL RECORD SHEET

Trainee P21 Department Hillsboro PD

Course Intro to Management Today's Date March 4, 1981

Have Action Plan? yes Interviewer Jones

ACTION ITEM:

Try to institute "vertical slice" groups for departmental activities, such as the purchase of police cars and other new equipment

BEHAVIOR: (# 1 of 2 related to this action item)

Discussed with chief the possibility of instituting vertical slice groups in the department

OUTCOME/RESULT:

The chief agreed with the idea and authorized me to establish a trial "vertical slice" committee of 5 officers to review the department's duty and procedures manuals.

JUDGMENT OF IMPACT:

+

PROBLEMS/CONCERNS:

A continuing labor dispute has hindered good relations between management and line personnel, creating suspicion about the purposes of a vertical slice committee.

EXHIBIT 39

NEP/Police Management Training

NONBEHAVIORAL RECORD SHEET

Trainee P21

KNOWLEDGE OR ATTITUDE CHANGES:

- Made me sensitive to the need for more line participation in department and unit decision-making
- Showed me the importance of the situational appropriateness of management styles

PLANNED BEHAVIORS:

- Develop a year-long schedule that meets documented training priorities
- Present the idea of establishing a crime prevention unit

WAYS COURSE SHOULD BE CHANGED TO BE MORE USEFUL:

Give more guidance in exercises and modify them so relate better to the needs of smaller departments

b. Develop system for counting and weighting behaviors. The behaviors reported by trainees are often not comparable in scope and in probable impact. It would be inappropriate to put a behavior such as "spoke to the chief about need for updated performance appraisal system" on a par with "successfully implemented MBO in my unit as a system for appraising performance." The second behavior obviously entailed multiple behaviors, even if they are not explicit. Had the interviewer succeeded in the probing task, then the sequence of distinguishable behaviors that consummated in implementation of a new performance appraisal system would be clearer. It is extremely difficult, after the fact, to try to separate a complex behavior into a set of behaviors, and we do not recommend trying to compensate for ineffective probing by recounting behaviors.

As an alternative, it may be feasible to develop a weighting system. This would involve assigning a value or weight to each behavior to reflect the complexity or magnitude of the effort. In our example, talking to the chief might get one point, but full implementation four points. We stress that this procedure is merely a way to compensate for prior interviewing drawbacks; we do not recommend its use unless program decision-makers require this type of information. Even if recorded behaviors consequently differ widely in their complexity, this will be self-evident to the program's decision-makers. Thus, they can draw conclusions from the qualitative data without putting much store in the quantitative.

c. List nonbehavioral data. Data from nonbehavioral record sheets for all trainees should be transferred onto one of three lists. The lists correspond to the three nonbehavioral information categories--specifically, knowledge or attitude changes, planned behaviors, and ways the course should be changed. These lists later provide the basis for answering questions about the course's nonbehavioral effects.

d. Develop individual trainee behavioral matrices. Transferring data from behavioral record sheets to individual trainee matrices is the final stage of coding. It is an important step because the resultant matrices constitute the primary reference document for most analyses that follow. In fact, in a straightforward evaluation that does not address the relative effects of several course aspects, the answers to most evaluation questions can be extracted directly from the matrices.

Exhibit 40 illustrates the type of information that individual trainee matrices should contain. At a minimum, they should have the information categories shown: a trainee code number, action items, behaviors, outcomes, problems, and trainee's judgment of impact. They may also contain other types of information (e.g., whether action items and/or behaviors were related to course objectives). As the exhibit shows, the matrices tend to fan out in the middle categories. For each trainee code in column 1, there may be several action items in column 2. For each action item in column 2, there may be several related behaviors in column 3. For nearly every behavior, there is at least one outcome in column 4, and often more than one. For a percentage of behaviors, ranging widely from trainee to trainee, there is one problem or more in column 5. Finally, column 6 simply places a value on the outcomes reported in column 4 by showing with a check mark whether the input was judged to be positive, negative, uncertain, or no change.

EXHIBIT 40(1)

NEP/Police Management Training

INDIVIDUAL TRAINING MATRICES

Participant	Action Items	Behaviors	Outcomes	Problems	Judgment Of Impact			
					+	-	+/-	NC
1	Prepare a reorganization plan and organization chart of the department.	Discussed with the Chief how POLEX materials could be applied in revising a 1979 departmental reorganization plan	The Chief was receptive to my ideas and authorized me to revise the plan and submit it to him for approval		x			
		Established a study group for revising the plan that consisted of the department's four "managerial" lieutenants	The study group was established but had to work with less rank-and-file input than there was in 1979, when the first reorganization plan was drawn up	The tense atmosphere in the department, resulting from a long-standing labor dispute, made it undesirable to obtain lower-level input in revising the reorganization plan			x	
		Reviewed the reorganization plan in a study group meeting that used POLEX organization principles, two prior consultant studies, and input from the Chief and other sergeants and lieutenants in the department	The study group meeting reviewed the plan and gave me several ideas for revision		x			
		Developed a new reorganization plan, based on the results of the study group meeting	An improved reorganization plan document and chart were developed, to be submitted to the Chief for review and approval	Getting enough time to complete the new plan, due to other job requirements	x			
		Submitted the new reorganization plan to the Chief for review and approval	Unknown					x
	Prepare a new system for formal communications within the department	Introduced AVO (Avoid Verbal Orders) forms within the department as an interim step until a new system for written communications can be developed	The use of AVOs got more people to write down their communications but was also unsystematic and time consuming, if done thoroughly; it didn't solve any of our major problems but didn't make the situation worse, either	Getting enough time even to start on this project, due to my regular duties and the special reorganization project			x	
		After discussion with another lieutenant, who had also attended POLEX and had listed developing a written communication system on his Action Plan, we agreed that he would develop the plan	I have a more manageable workload and have not taken on too many tasks to complete successfully	Not enough time really to start developing the system, due to job responsibilities and the reorganization plan	x			
Subtotal	2	7	7	4	4	0	2	1
		I talk more with the other lieutenants in the department, since I had to work with some of them in POLEX	I'm more aware of the value of talking and working with others I feel more involved in the department		x			
Subtotal	0	1	2	0	1	0	0	0
Total	2	8	9	4	5	0	2	1

EXHIBIT 40(2)

Participant	Action Items	Behaviors	Outcomes	Problems	Judgment Of Impact			
					+	-	+/-	NC
2	Enact a system for setting up, one year in advance, a series of training meetings, with each meeting to be based on training need information elicited from the entire department	Discussed with the Chief the idea of establishing a system for identifying departmental training needs through input from the entire department and then developing and scheduling training sessions to meet those needs	The Chief agreed with the idea and told me to go ahead with it		X			
		Organized my ideas into a letter to be distributed throughout the department eliciting opinions and suggestions on current and future departmental training needs	Developed a draft letter to send throughout the department to identify training needs		X			
		Submitted the letter to the Chief for approval	The Chief approved the letter		X			
		Distributed letter throughout the department	Two responses have been received to the letter; both have constructive and favorable		X			
	Recommend to the Chief that the Profile of Organizational Characteristics be administered to everyone in the department	Discussed with the Chief the possibility of administering the the Profile of Organizational Characteristics to everyone in the department	Although the Chief approved of of the questionnaire, he said it would be too costly and time consuming to administer throughout the department	The department's budget is very tight. As a result, the department has been shorthanded and everyone has been very busy			X	
Total	2	5	5	1	4	0	1	0
3	Develop a memoranda system to be used by all employees of the department in written communications within the department	Discussed with the Chief the inadequacies of current written communications procedures and the need for a new system to correct those inadequacies	The Chief agreed with me and said I should develop the new system by March 1	I am the crime prevention officer in the department and my duties peak in December. Because of this I wasn't able to start on the memo system until mid-January	X			
		Made phone calls to seven local departments to find out about their memoranda systems and obtain copies of their memo forms	Three of the seven departments called employed a formal memoranda system. These three gave me useful information and sent me copies of their forms		X			
Subtotal	1	2	2	1	2	0	0	0
		I asked a sergeant and a lieutenant in the department for suggestions on how to distribute crime prevention articles to as many citizens as possible	I increased my distribution of the articles by one-half after taking their suggestions to distribute them to the Boy Scouts and the Knights of Columbus		X			
Subtotal	0	1	1	0	1	0	0	0
Total	1	3	3	1	3	0	0	0

EXHIBIT 41

NEP/Police Management Training

OUTCOMES OF TRAINEE ATTEMPTS
AT IMPLEMENTING NEW BEHAVIORS
(POLEX)

Type Of Outcome	# Of Times Reported	# Of Trainees
Outcomes Related To Chief, Superior, And Subordinate Reactions To Implementation		
1. Chief thought attendance at course was worthwhile	3	3
2. Chief agreed with suggestions and recommendations and authorized implementation activities	22	14
3. Chief agreed with suggestions but refused to authorize present implementation	10	5
4. Chief disagreed with suggestions and recommendations and refused to authorize implementation	6	3
5. Superior received and approved recommendations and agreed to pass them on to the Chief	1	1
6. Prepared presentations to Borough Council and subcommittees	2	2
7. Borough Council Members agreed with the recommendations and authorized presentation to the full Borough Council	1	1
8. Subordinates understood and were receptive to suggestions	6	5
Outcomes Related To Information Collection And Committee Meetings		
1. Useful information was obtained for implementation through personal information collection activities	15	6
2. Useful information was obtained for implementation through committee and group meetings	6	3
3. Useful information was obtained for implementation through completed surveys and forms	3	3
4. Completed surveys and forms indicated that the men misunderstood the forms, refused to cooperate, or did not take the forms seriously	1	1
Outcomes Related To Implementation Products And Decisions		
1. Further implementation facilitated	7	7
2. New and revised plans, manuals, and forms were developed	6	6
3. Committees were established	6	4
4. Committees decided on new implementation objectives	2	1
5. Decided that further activities were undesirable	4	4
Outcomes Related To Implementation Results		
1. Personal schedule and information are better organized	2	2
2. Personal performance results have improved	5	4
3. Personal morale has improved	4	3
4. Unit activities and information are better organized	5	3
5. Unit results have improved	3	2
6. Unit morale has improved	7	5
7. Implementation activities have not solved major problems	1	1
Unknown Outcomes	7	6
Total	135	17

EXHIBIT 42

NEP/Police Management Training

PROBLEMS ENCOUNTERED BY TRAINEES
IN IMPLEMENTING NEW BEHAVIORS
(POLEX)

Problems	Number Of Times Reported	Number Of Trainees
Lack Of Time And Resources For Implementation		
1. Getting enough time away from regular duties	15	6
2. Getting enough time, due to seasonal short-handedness	2	2
3. Finding the most convenient time for meetings, due to conflicting duty schedules	5	2
4. Lack of resources due to budget constraints	2	2
Inability To Convince Superiors And Subordinates To Cooperate In Implementation		
1. Chief is authoritarian and resistant to change	4	3
2. Subordinates refuse to cooperate in implementation	5	5
3. Subordinates misunderstand the details of implementation	2	2
4. Lack of support on Borough Council	1	1
Miscellaneous Problems		
1. Decisions were postponed while Chief recovered from illness	3	2
2. I lacked the necessary knowledge of MBO to push for its implementation	1	1
3. A patrolman hampered an ongoing investigation by divulging materials in a case report to suspects in the investigation	1	1
4. Department moved to new facilities	1	1
5. Subordinates forgot about implementation	2	2
Total	44	15

EXHIBIT 13

NLP/Police Management Training

NEW BEHAVIORS REPORTED BY TRAINEES (POIEX)

1	2	3	4	5	6	7		8	
Participant	Number of Action Items	Number of Behaviors Related to Action Items	Number of Other Behaviors	Total New Behaviors	Percentage of Action Items Leading to New Behaviors	Number of Behaviors Related to Action Items/ Number of Action Items		Number of Behaviors Related to Action Items/Total Number of Behaviors	
1	1	2	4	6	100%	(2/1)	2:1	(2/6)	1:3
2	2	7	1	8	100	(7/2)	3.5:1	(7/8)	7:8
3	2	4	3	7	50	(4/2)	2:1	(4/7)	4:7
4	2	6	0	6	100	(6/2)	3:1	(6/6)	1:1
5	5	15	4	19	100	(15/5)	3:1	(15/19)	15:19
6	1	1	1	2	100	(1/1)	1:1	(1/2)	1:2
7	3	3	2	5	100	(3/3)	1:1	(3/5)	3:5
8	1	4	0	4	75	(4/4)	1:1	(4/9)	1:1
9	2	5	0	5	100	(5/2)	5:2	(5/5)	1:1
10	3	3	0	3	100	(3/3)	1:1	(3/3)	1:1
11	3	8	2	10	100	(8/3)	8:3	(8/10)	4:5
12	2	2	5	7	[illegible]	(2/2)	1:1	(2/7)	2:7
13	3	1	4	5	33	(1/3)	1:3	(1/4)	1:4
14	1	2	1	3	100	(2/1)	2:1	(2/3)	2:3
15	2	4	1	5	100	(4/2)	2:1	(4/5)	4:5
16	4	7	0	7	100	(7/4)	7:4	(7/7)	1:1
17	1	4	1	5	100	(4/1)	4:1	(4/5)	4:5
TOTAL	41	78	29	107	90%	(78/41)	2:1	(78/107)	78:107

Questions about the overall extent of course effects--in terms of action items, planned and unplanned behaviors, problems, and classification of impacts--may be answered by performing simple manipulations on the quantitative data in the matrices. We illustrate below how we answered a few of the most commonly asked questions in our evaluation of a POLEX course. To what extent did trainee plans lead to new behaviors? To what extent did the course result in behaviors that were unplanned, in that they were not expressed as end-of-course intentions? To what extent were changes in trainee behavior planned? Exhibit 43 shows how data were arrayed to answer these questions. The table summarizes in columns 1 through 5 information that was drawn directly from individual trainee matrices: participant codes, number of action items, number of behaviors related to action items, number of other behaviors, and total new behaviors. The information for column 6, the percentage of action items leading to new behaviors, was obtained by eyeballing the trainee matrices. Columns 7 and 8 were filled by taking ratios of information contained in columns 3 and 2 (for 7) and 3 and 5 (for 8). So, to what extent did trainee plans lead to new behaviors? In Exhibit 43, we can see the number of planned behaviors, the ratio of planned behaviors to action items, and the percentage of action items resulting in at least one new behavior. To what extent did the course result in unplanned behaviors? We can derive from Exhibit 43 the number of trainees who reported unplanned new behaviors, the average number of unplanned behaviors reported, and the distribution of new behaviors (i.e., whether a few trainees accounted for most of them). To what extent were new behaviors planned rather than unplanned? Column 8 provides an answer to this question.

To what extent did trainees encounter problems in implementing new behaviors? Exhibit 44 shows how this question was answered, relying entirely on the individual trainee matrices and then calculating simple ratios. From this, we can see the number of problems encountered, the number of trainees with problems, and the average number of problems per trainee.

Other questions can similarly be answered primarily from data in the matrices. To determine whether action plans or resultant behaviors generally reflected course objectives, however, it would first be necessary to have a set of objectives that the program accepts and against which plans and behaviors may be judged. Once plans and/or behaviors have been compared against objectives and each checked off as reflecting or departing from course objectives, simple ratios may be calculated.

d. Examine training's specific effects. Once relevant questions about training's overall effects have been satisfactorily answered--by visually scanning the individual trainee matrices and/or by constructing new classifications of qualitative and quantitative data--training's specific effects may be examined. Here, the evaluator is trying to determine the effects of different aspects or parts of a course. It is important to know how the different parts of a course affect job behavior so that the course can be effectively modified to maximize transferability. There are three frames of reference within which the evaluator might choose to examine a course's differential effects. If course objectives are sufficiently clear, the evaluator might examine course effects against each objective, in this way determining the extent to which the course supports its stated objectives. In the typical course, however, objectives will not be clear enough to do this, so one of two alternatives may be

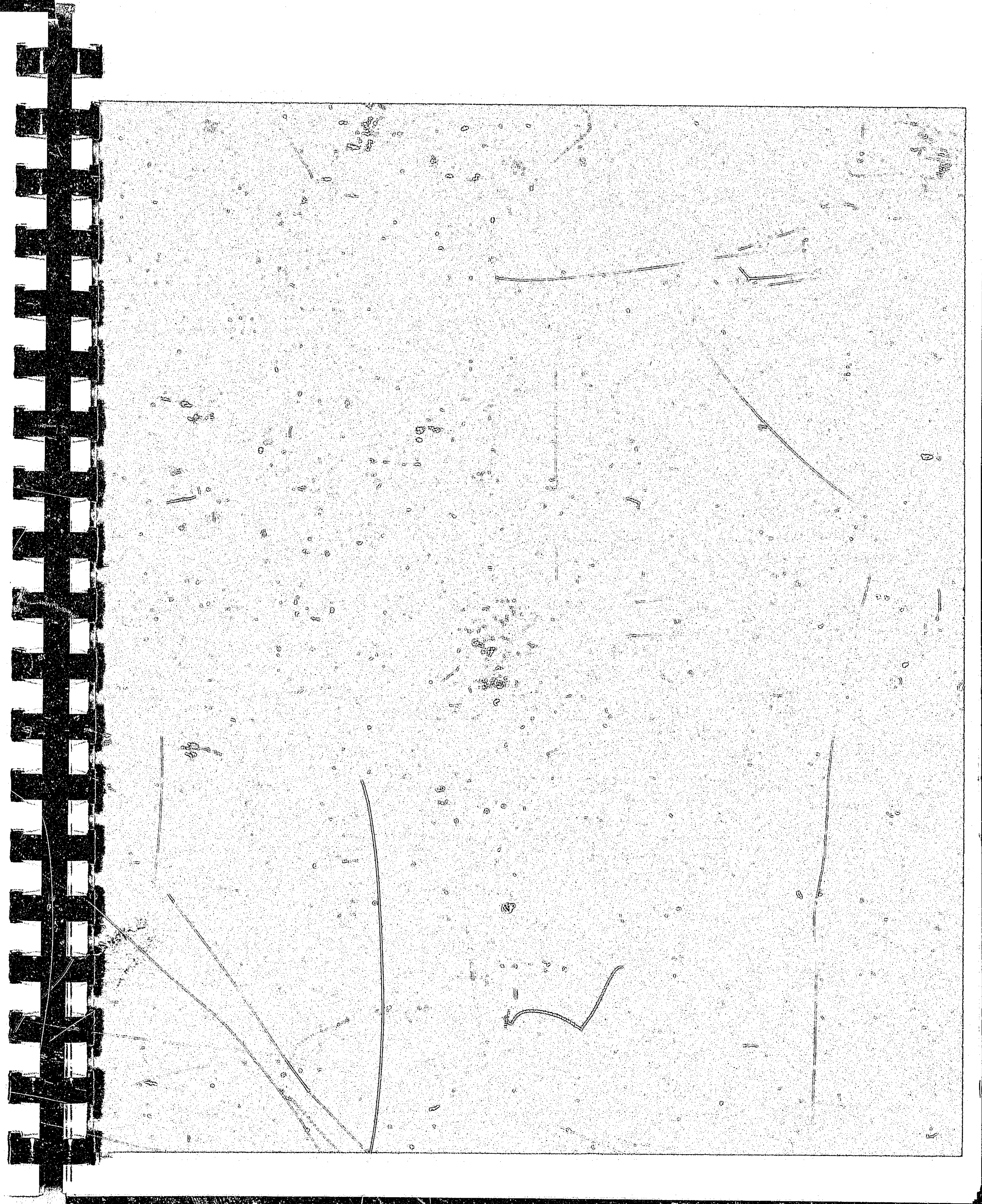


EXHIBIT 43

NEP/Police Management Training

NEW BEHAVIORS REPORTED BY TRAINEES
(POIEX)

1	2	3	4	5	6	7	8
Participant	Number of Action Items	Number of Behaviors Related to Action Items	Number of Other Behaviors	Total New Behaviors	Percentage of Action Items Leading to New Behaviors	Number of Behaviors Related to Action Items/ Number of Action Items	Number of Behaviors Related to Action Items/Total Number of Behaviors
1	1	2	4	6	100%	(2/1) 2:1	(2/6) 1:3
2	2	7	1	8	100	(7/2) 3.5:1	(7/8) 7:8
3	2	4	3	7	50	(4/2) 2:1	(4/7) 4:7
4	2	6	0	6	100	(6/2) 3:1	(6/6) 1:1
5	5	15	4	19	100	(15/5) 3:1	(15/19) 15:19
6	1	1	1	2	100	(1/1) 1:1	(1/2) 1:2
7	3	3	2	5	100	(3/3) 1:1	(3/5) 3:5
8	4	4	0	4	75	(4/4) 1:1	(4/9) 1:1
9	2	5	0	5	100	(5/2) 5:2	(5/5) 1:1
10	3	3	0	3	100	(3/3) 1:1	(3/3) 1:1
11	3	8	2	10	100	(8/3) 8:3	(8/10) 4:5
12	2	2	5	7	100	(2/2) 1:1	(2/7) 2:7
13	3	1	4	5	33	(1/3) 1:3	(1/4) 1:4
14	1	2	1	3	100	(2/1) 2:1	(2/3) 2:3
15	2	4	1	5	100	(4/2) 2:1	(4/5) 4:5
16	4	7	0	7	100	(7/4) 7:4	(7/7) 1:1
17	1	4	1	5	100	(4/1) 4:1	(4/5) 4:5
TOTAL	41	78	29	107	90%	(78/41) 2:1	(78/107) 78:107

EXHIBIT 44

NEP/Police Management Training

EXTENT OF PROBLEMS ENCOUNTERED BY TRAINEES IN IMPLEMENTING NEW BEHAVIORS (POLEX)

Trainee Number	Number Of New Behaviors	Number Of Problems	Ratio Behaviors/ Problems
1	6	6	1:1
2	8	4	2:1
3	7	0	--
4	6	0	--
5	19	9	19:9
6	2	3	2:3
7	5	3	5:3
8	4	1	4:1
9	5	1	5:1
10	3	3	1:1
11	10	2	5:1
12	7	6	7:6
13	5	2	5:2
14	3	1	3:1
15	5	2	5:2
16	7	1	7:1
17	5	1	5:1
Totals	107	44	5:2

88.8 percent of trainees encountered at least one problem in implementing behaviors.

11.1 percent of trainees encountered no problems in implementing behaviors.

tried. Course effects may be examined module by module, or, if this too seems unrealistic, behaviors may be grouped naturally and other variables of interest examined in the context of the natural behavioral groupings.

Any of the variables germane to the overall analysis may be focused upon in the specific analysis. It might be important, for example, to determine exactly which course objectives (or other course aspects) are most prominent in action plans. Similarly, program decision-makers might need to know which aspects of a course were most clearly reflected in changed trainee behavior, which yielded the greatest number of positive impacts upon trainees' agencies, and which tended to produce implementation problems.

The first step is to select a frame of reference. Let us choose natural behavioral groupings for illustrative purposes. Exhibit 45 shows how we grouped behaviors for the POLEX evaluation. To develop this classification scheme, we reviewed the behavioral record sheets, drew up a preliminary set of categories, sorted the record sheets into these categories, and, by re-sorting the record sheets several times, came up with an acceptable scheme.

From this Exhibit, we can draw conclusions about aspects of the course that had the strongest influence. Because the purpose of later analyses was to see how other variables looked in this behavioral framework, we had to label each behavioral record sheet to indicate the type of behavior involved. For example, any record sheet that reflected revision or reorganization of a specific departmental activity was marked (I). If this specifically involved development of a departmental reorganization plan, it was also labeled (1); revisions of departmental written communications systems were labeled (2), and so forth, in line with Exhibit 45. These labels helped to sort out data in subsequent analyses. Exhibit 46, for example, examines trainee judgments of the impacts of their behaviors within the natural behavioral grouping. Similar analyses may be done for other variables, including outcomes and problems, by type.

e. <u>Examine training's nonbehavioral effects</u>. It is sometimes also important to examine training's nonbehavioral effects. We include in this category knowledge and attitudes that have not affected behavior, planned behaviors that have yet to be implemented, and trainee recommendations for course improvement. These can be qualitatively and quantitatively meaningful, if seen in relation to behavioral data. If nonbehavioral change is of the same general magnitude as behavioral change, this tends to reinforce conclusions about the course's influence on behavior. If nonbehavioral change is of a lesser magnitude than behavioral change, then the course's transferability would seem to be supported. However, if nonbehavioral change is of greater magnitude than behavioral change, then we know that something is wrong. Either we followed up too soon, or the training does not focus adequately on transferability, or trainees' agencies were resistant to change, or the wrong trainees were sent to training, or evaluators did an ineffective job of probing, or something else went wrong. Similarly, if nonbehavioral changes differ from behavioral change in their focus, this can tell us something about areas of weakness. Trainee recommendations for course improvement can also be viewed in contrast to other data.

EXHIBIT 45

NEP/Police Management Training

NATURAL GROUPINGS OF NEW
REPORTED BEHAVIORS
(POLEX)

Types Of Behaviors	Number Of Behaviors Reported	Number Of Trainees Reporting
Revised Or Reorganized Specific Departmental Activities		
1. Developed a departmental reorganization plan	8	3
2. Revised the department's written communications system	14	5
3. Introduced the "concurrence/non-concurrence"directive form	8	2
4. Revised case reporting forms and procedures	10	2
5. Changed unit property receipt forms	1	1
6. Changed unit evidence collection procedures	1	1
Promoted Increased Participation In Departmental Decision-Making		
1. Established vertical slice groups	10	5
2. Attempted to implement MBO in the department	7	5
3. Urged the Chief to delegate more authority	1	1
4. Promoted overall participation in department and unit activities	6	5
Prepared And Circulated Information To Superiors, Co-Workers, And Subordinates		
1. Developed a unit information flyer	2	1
2. Held weekly information and discussion meetings in the division	1	1
3. Discussed POLEX and circulated POLEX information to the Chief and others in the department	6	4
Improved Subordinate Attitudes And Performance		
1. Developed improved training programs	9	2
2. Attempted administration of the Profile of Organizational Characteristics	9	5
3. Applied course principles in appraising subordinates	7	3
4. Acted more "hard-nosed" to subordinates, as necessary.	1	1
Communicated More Frequently With Co-Workers		
1. Talked informally with co-workers and subordinates	2	2
2. Solicited advice from co-workers on performance of duties	1	1
Introduced Improved Police Practices In The Department		
1. Attempted to implement safer speed zones	2	1
2. Attempted to change Borough road signs to safer "break-away" signs	4	1
3. Attempted to institute a crime prevention unit in the department	4	1
Total	114	17

EXHIBIT 46

NEP/Police Management Training

TRAINEE JUDGMENTS OF IMPACT
FOR REPORTED BEHAVIORS
(POLEX)

	Reported Behaviors	Judgments Of Impact			
		+	-	±	NC
I.1.	Developed A Departmental Reorganization Plan	4	2	1	2
I.2.	Revised The Department's Written Communications System	10	0	0	2
I.3.	Introduced The "Concurrence/Non-Concurrence" Form In The Department	6	0	1	0
I.4.	Revised Case Reporting Forms And Procedures In The Department	8	0	0	2
I.5.	Revised Unit Property Receipt Forms	1	0	0	0
I.6.	Revised Unit Evidence Collection Procedures	1	0	0	0
II.1.	Established Vertical Slice Groups Within The Department	1	1	0	0
II.2.	Attempted To Implement MBO In The Department	4	0	2	0
II.3.	Urged The Chief To Delegate More Authority	1	0	0	0
II.4.	Promoted Increased Participation In The Department	3	1	0	0
III.1.	Developed A Unit Information Flyer	2	0	0	0
III.2.	Held Weekly Division Information And Discussion Meetings	1	0	0	0
III.3.	Discussed And Circulated Course Information With The Chief And Co-Workers	4	0	0	0
IV.1.	Developed Improved Training Programs	7	0	0	2
IV.2.	Administered The Profile Of Organizational Characteristics Within The Department	2	1	2	2
IV.3.	Applied Course Materials To The Appraisal Of Subordinates	5	1	0	0
IV.4.	Acted More "Hard-Nosed" To Subordinates, As Necessary	0	0	0	1
V.1.	Talked Informally With Co-Workers	1	0	0	0
V.2.	Solicited Advice From Co-Workers On Performance Of Duties	1	0	0	1
VI.1.	Attempted To Implement Lower Speed Zones	2	0	0	0
VI.2.	Began Changing Borough Road Signs To Safer "Breakaway" Signs	4	0	0	0
VI.3.	Attempted To Institute A Crime Prevention Unit In The Department	3	0	0	1
	Total	71	6	6	12

The analyses of training's nonbehavioral effects are simple to do. During the coding process, three lists of nonbehavioral effects are developed from the trainees' nonbehavioral record sheets. If the evaluator is only interested in estimating the relative magnitude of behavioral and nonbehavioral change, these lists do not ever have to be thoroughly read. The number of entries alone will give that information. To determine more precisely the nature of non-behavioral change, the difference in thrust between behavioral and nonbehavioral change, and recommended course changes, the evaluator may construct natural groupings from the three lists. Exhibit 47 shows how we grouped the attitudinal and knowledge changes reported by POLEX trainees. The relative magnitude of certain behavioral and nonbehavioral changes can also be examined more precisely by going back to the nonbehavioral record sheets and extracting, trainee by trainee, the number of nonbehavioral changes of the two main types and by then calculating ratios, thus constructing a matrix similar to Exhibit 43.

8. Assess Implications In Evaluation Report. Once the evaluator has constructed and reviewed the matrices related to the evaluation questions, the Evaluation Report may be written. In doing so, several considerations should be borne in mind.

a. Include introductory information that provides shortcuts for the reader. Because most readers read selectively, to their own interests, it is important to help the reader find information relevant to his or her interests as early as possible. The front of the report should include, therefore, a list of exhibits, a list of evaluation questions (with corresponding page references), and a summary of the evaluation's background and processes.

b. Provide an overview of conclusions and recommendations. Although it is proper to provide shortcuts to the reader who reads to selective interests, is also important that each reader have a balanced overview of the evaluation's conclusions and recommendations. Accordingly, a one- or two-page summary of this sort should be placed at the front of the report, before the indices.

c. Develop a separate section to answer each evaluation question. Each evaluation question should be addressed in a separate narrative section, even if this means some sections are less than a page in length. The discussion should refer to related matrices, summarize the data in the matrices, and concisely explain how the data answer the question.

d. Address implications for the course. In the concluding section of the report, the implications of evaluation data for the course should be discussed. Because most of the data gained by the approach may be interpreted in several different ways, drawing implications requires the evaluator to weigh alternate judgments carefully. For any evidence of discrepancy between intentions and implementation, the evaluator may draw several conclusions. Training might not have provided trainees with enough information to carry out their plans. Trainees might not have been sufficiently motivated by training. The time needed for implementation might need to be longer than that allowed before the follow-up. The agency context may have been an inappropriate trying ground for trainee plans. In reaching such conclusions, the specific analyses should point out areas of the course to which the conclusions are especially relevant.

EXHIBIT 47

NEP/Police Management Training

KNOWLEDGE AND ATTITUDE CHANGE (POLEX)

Key Types Of Nonbehavioral Change	Number Of Changes Reported	Number Of Trainees Reporting
Course Allowed Trainees To See A Management Perspective On Departmental Activities	7	7
Course Made Trainees Sensitive To The Need For More Information In Order To Accomplish Projects Most Effectively	6	6
Course Made Trainees Sensitive To The Need For Participation In Department And Unit Decision-Making	6	6
Course Made Trainees More Rigorous In Their Thinking On The Job	6	6
Course Made Trainees Sensitive To The Situational Appropriateness Of Management Styles And Techniques	3	3
Course Reinforced Information Obtained In Past Experience Or In Other Police Management Training Programs	3	3
Course Made Trainees Aware Of The Need For Working With Other Officers	5	5
Course Made Trainees' Attitudes Toward Policing More Positive	3	3
Course Stimulated Trainees To Pursue Advanced Education	1	1
Course Frustrated Trainee, Due To Resistance To Change In His Department	1	1
Course Improved Overall Trainee Proficiency On The Job	1	1
Total	43	17

The evaluator's ultimate concern should be the identification of options for adjustment of program expectations, activities and information systems. Depending on how convincingly the evidence points to a particular set of conclusions, the evaluator may need to consider a range of potential program adjustments. The evaluator might recommend that the program collect more data to pinpoint the sources of difficulty; follow up on behavior again at a later point in time; clarify the program's objectives to the sending organizations; try to shift to a different training audience or, at least, include a broader mix of ranks; send memos to trainees' superiors asking for assistance in implementation; ask superiors to work with trainees to develop personalized pretraining objectives; discuss implementation problems at greater length in training; drop certain modules and replace them with others; beef up the substantive information and supporting exercises in certain modules; and so forth. It is essential that the evaluator work closely with program managers in identifying action implications for the program before recording them in the evaluation report because this helps ensure buy-in and eventual consideration of implications.

B. PROFICIENCY ANALYSIS APPROACH

The Proficiency Analysis (PA) Approach--which OPM calls "the needs part of Value Model II"--examines the question: Does training improve the trainee's job proficiency in ways that the organization can use? This approach can help answer several specific questions, such as:

- At what proficiency levels are trainees required to perform on their jobs?
- At what proficiency levels do trainees actually perform?
- What managerial training needs do trainees have and how great are they?
- Do trainees and others who work with them, especially supervisors, agree on how to prioritize the trainees' training needs?
- Do trainees think that their training priorities have changed favorably as a result of training, i.e., that their job proficiency has improved?
- Do supervisors think that trainee job proficiency has improved as a result of training?
- How much do trainees and supervisors agree on the areas where trainee job proficiency has improved?
- In what knowledge and ability areas should the course focus more attention to be more job-related?

The underlying concept behind PA is that successful training ought to decrease perceived training needs, i.e., it should close gaps between reported job requirements and current capability levels. Whether and how training succeeds in closing performance gaps (and, hence, in increasing proficiency) can be

determined by comparing training needs reported at two points in time: before training and a few months after the trainee has returned from training to the job. It can also be very useful to obtain more than one perspective on changes in training needs, as from a supervisor or other co-worker.

The problem that most public agencies have faced in carrying out similar approaches is that few have resources to conduct a task analysis of managerial positions. Without a managerial task analysis, it is difficult to identify needs in terms of explicit job requirements. Recognizing this problem, OPM constructed a standard instrument called the Managerial Training Needs Profile (MTNP). It measures 14 generic managerial functions or activities (ones that all managers are required to possess, to varied degrees). It also measures generic knowledges and abilities in which management training might be needed. On each of the 14 activities, it has managers (or their co-workers) rate the minimum required proficiency level for the position, the manager's current capability level, and an importance weight. These three ratings are factored together to prioritize training needs among the 14 activity areas. On each of the knowledges and abilities, it has respondents rate the minimum required level for the position and the manager's current capability level. These ratings are used to identify the knowledges and abilities in which an individual manager, or a group of managers, require training and the proficiency levels to which they need to be raised.

Certain preconditions for and cautions about the use of PA should be understood. The OPM authors stress that this approach is especially suited to evaluate programs that aim at behavioral changes not amenable to direct observation. They also caution, however, that it should <u>not</u> be used when proficiency can be measured more directly. We add that it also should not be used if the course is of such duration or nature that measurable change in proficiency is implausible. In addition, if the evaluation has to permit tests of statistical significance, PA either should not be used or several classes from the same course should be combined for the analysis. In some instances, PA may be appropriate but there may be problems with using the MTNP as the proficiency measure. For example, it is inappropriate to use the MTNP as the framework for measuring proficiency when a good task analysis of the study sample's managerial functions is available. The MTNP should also not be used if trainee backgrounds and educational levels are such that trainees would resist the instrument or find it hard to interpret. Because PA allows for a posttraining re-acclimation period, it can only be used when the evaluator is allowed several months to produce evidence of course effectiveness. PA requires certain staff skills as well. The individual who introduces the PA concept has to have sufficient credibility with the target group to persuade them of its usefulness and assuage fears that data will be misused (e.g., in a performance appraisal). This is especially important when trainees know that superior ratings will also be obtained. The individual who analyzes PA data needs moderately strong analytical skills and may also need programming skills, if the analyses to be done are not included in OPM's MTNP computer package.[14] PA also requires computer access. But it entails no major adjustments

---

14/ Sample copies of the MTNP and of the MTNP/Value Model II computer package may be obtained from OPM's Training Consulting Division at the address given in footnote 9.

in course orientation and limited alteration of structure, unlike APFUA. The exact nature of the needed accommodation depends on when and how the MNTP is to be administered. If trainees are supposed to complete it before they arrive at the training facility, then it should be included in pretraining materials; if they are to fill it out after they arrive, then it has to be incorporated into the course curriculum as an exercise or at least introduced within a course module. Beyond that, no specific adjustments are necessary.[15]

This PA overview contains five major steps, each with several substeps. This description should be adequate to permit use of PA, but, for supplementary information, the following resources will be helpful:

- OPM pamphlets on this approach and related approaches[16]
- Related articles in Peterson's anthology on management training evaluation and in other sources[17]
- Opinions of the training program director who hosted our feasibility test[18]
- Assistance from OPM's Training Consulting Division (in Washington, D.C.) or from an OPM regional office

1. Resolve Preliminary Evaluation Questions. Several preliminary questions should be answered before attempting to use this approach.

---

[15] There are precedents for this approach to course evaluation, but it is not MTNP's conventional use. Typically, the MTNP is administered only before training. It is used in one of two ways. Group profiles provide the basis for course design; individual profiles help to select and assign managers to appropriate training courses.

[16] These include: S.S. Rittenhouse, A.L. Breitler, and R.G. Phillips, Needs Assessment and Evaluation of Training. (Washington, D.C.: Office of Personnel Management, April 1980); Report of Managerial Training Needs for Sample Application (Washington, D.C.; Office of Personnel Management, July 1980). Both may be obtained from OPM's Workforce Effectiveness and Development Group, Office of Consulting Services, Training Consulting Division.

[17] Peterson (ed.), op. cit., contains an article on Simplified Cost-Benefit Analysis that encompasses PA. See G.R. Seppala, "An Approach to Determining the Value of Management Training," pp. 180-202. See also D.L. Poster, A.L. Breitler, and S.S. Rittenhouse, "NET: A Methodology for Needs Assessment and Evaluation of Training" (tentative title), "Performance, 1(6), forthcoming.

[18] The director may be reached at: Northwestern University, The Traffic Institute, 555 Clark Street, Evanston, IL 60204.

a. Determine what questions the evaluation will address. For which questions related to the approach do the program's decision-makers want useful information? It is helpful but inessential to agree entirely on the evaluation questions before data collection. In fact, certain evaluation questions tend to arise only after data analysis has begun, when initial analyses do not fit expectations. Still, it is valuable to establish early what questions program decision-makers are really concerned with, so analyses may be planned and the evaluation's purposes legitimately represented.

b. Assess the relationship between PA and other evaluative approaches. Does PA tie into existing evaluation approaches? Are there more direct ways to obtain the information it will generate? Can it really answer questions that the program should be trying to answer? The evaluator should deal with questions like these to ensure that PA is an appropriate evaluation technique.

c. Decide whose viewpoints on training needs will be obtained. Would a second viewpoint on training needs be useful? When a perspective other than the trainee's is obtained, it is typically the supervisor's. The main advantage of deciding in favor of a second perspective is that it can bolster or validate the pretraining profile of job requirements. In other words, it can give assurance that trainees and supervisors have compatible expectations about the level at which trainees ought to perform. With the second perspective, one can also get a better sense of how effectively a course is meeting the organization's needs. There are also disadvantages to a second perspective, however. First, supervisory data tend to be much harder to collect than trainee data because supervisors almost always have to be contacted exclusively by mail. Second, supervisory views of changes in proficiency tend to understate the degree of change, especially if follow-up data are collected before supervisors have had sufficient opportunity to observe changes directly. Third, if trainees know about supervisor involvement, they may become suspicious that data will be misused and, consequently, decline to give full cooperation. These considerations should be weighed carefully, but we assume in later discussions that supervisors' perceptions will be tapped.

d. Determine the timing and means for pretraining data collection. When and how should pretraining data be collected? Validity is probably maximized if the MTNPs are completed before training, while the trainees are still on the job. The MTNP's designers typically come right into an agency and invite its managers to fill out the MTNP before training. But they are rarely faced with later collecting posttraining data, they apply PA in one agency at a time, and they infrequently deal with courses attended by multiple agencies. Given the circumstance of working with a single police agency, it could be appropriate to come into the agency before training to have the MTNP completed.

In more typical circumstances, it may be difficult to collect data while prospective trainees are still on the job. Perhaps the best way to do so is to make completion of the MTNP a required part of course prework. It can be difficult to impose prework on the trainee's supervisor, however; requiring supervisor cooperation may cause some supervisors to send their officers elsewhere for training. If on-the-job pretraining data collection presents insurmountable problems, data collection may be deferred until trainees arrive for the course.

The best time to approach them with the MTNP is generally during the orientation period. Although trainee-instructor rapport tends to be low at this point, trainee perceptions of their job requirements and current performance levels have yet to be contaminated by exposure to the course. If trainee-instructor rapport is so low that data simply cannot be collected during the orientation period, then pretraining data should be collected as soon as possible thereafter. The longer one waits, the more perceptions become colored by the course. And at some point--fairly early in a course, we imagine--views can be so significantly altered that they no longer validly represent pretraining perceptions. Assuming that pretraining data collection is deferred until the start of a course, supervisory ratings would then have to be obtained in parallel with the course. This still leaves open the option of more creative data collection strategies, such as collecting data from supervisors in the application for course admission and later collecting trainee data during the orientation period. We assume in later discussions that trainee data will be obtained in an orientation exercise and supervisor data by mail while the course continues.

e. Determine the timing and means for follow-up data collection. When and how should posttraining data be collected? The means is rarely at issue. If trainees come from several widely dispersed agencies, there is no alternative but to collect follow-up data by mail. The timing of the follow-up still has to be resolved, however. As in APFUA, the follow-ups should not be made too soon, or trainees will not have had sufficient time to demonstrate their changed proficiency to themselves and to their supervisors; nor should they be put off too long because participants tend to lose interest and because, the longer the time interval, the greater the chances that other factors will intervene to change behavior, including promotions and transfers. The optimal follow-up period depends on the nature of the course and the immediacy of need for feedback. Follow-up should generally be conducted within two to three months after training.

f. Assign responsibilities for PA tasks. Who will have responsibility for introducing the PA concept, collecting pre- and posttraining data, analyzing results, and presenting them to the program's decision-makers? There is no inherent reason why one individual could not assume responsibility for all of these tasks. There is also no reason why an individual with an obvious stake in the evaluation's results cannot collect the data because, unlike APFUA, there is little or no need for direct evaluator-respondent contact. Typically, the course coordinator or lead instructor introduces the PA concept, explains its value for participants and the program's operators, and distributes the MTNP forms. The individual who assumes these introductory functions need not be directly involved in the course and, in many instances, will occupy a higher administrative function. This stature can add to the individual's credibility with the group. After initial distribution of the forms, a second individual might present a technical explanation of the instrument and respond to questions. After the MTNP has been completed, the course coordinator will typically collect identifying data and explain the general procedures for the follow-up. The collection of pretraining data from supervisors and from both groups after training requires little more than basic administrative capability and will typically be delegated to support personnel. The data analyses need not be done by an individual with course familiarity, but whoever presents results to the

program's decision-makers ought to have at least general knowledge of the course. It is essential to negotiate at least the in-classroom PA responsibilities well in advance.

g. Identify guarantees that will be given. Will certain guarantees have to be given for PA to be accepted? Especially when the evaluation is not done in-house, guarantees may be required. These can deal with issues such as administrative costs, confidentiality, and access to and permissible uses of results. The evaluator and course administrator should exchange guarantees in writing, and, where appropriate, guarantees should be shared with trainees and supervisors.

2. Collect Pretraining Proficiency Data. Once the preliminary questions have been resolved, pretraining data collection should be fairly straightforward.

a. Explain the evaluation process and its potential value. An individual with high credibility and natural rapport with trainees should explain the sources from which data will be collected, when and how this will be done, and the evaluation's potential value to respondents, to the program's operators, and to future trainees. It may be appropriate in some programs to defer explaining the supervisor's involvement until after trainee data have been collected. The tradeoff is that trainees may perceive that information was held back, become suspicious of the evaluation's "real purpose," and withdraw cooperation.

b. Present the MTNP as a course exercise. Trainees should be prepared for completing the MTNP just as they would for any other exercise. To the extent feasible, the MTNP exercise should be interwoven with other course activities. Adequate time should be given for all trainees to complete the MTNP--roughly 30 to 45 minutes. The MTNP forms should be divided into separable question-and-answer booklets so that trainees may place them side by side, thus reducing transposition errors. Each answer booklet should be precoded with an identifying number, which the trainee should be asked to record.

c. Allow time for review and correction of MTNP. After all trainees have had adequate time to complete the MTNP, time should allowed for trainee questions and for review of MTNPs for completeness. The review for completeness will occur while trainees are completing the identifying information sheet and the corrections afterwards.

d. Collect identifying information on a separate sheet. Only identifying information that will be required for the posttraining follow-up should be collected. The precoded form should, at a minimum, ask for the trainee's name, department, mailing address, and telephone number (in case the follow-up form is incompletely filled out). Supervisor's name, address, and telephone number should not be requested if these can be taken off the course admission application. Additional information may be collected, if directly relevant to the evaluation, such as rank, educational background, and age.

e. Explain mechanics of follow-up to trainees. Trainees need to know how they will receive the follow-up MTNP form (by mail), under whose name it will be sent, when they should expect it, and how long they will have to

return it. They should also be reminded that supervisors will receive a copy at the same time.

f. Send MTNPs to supervisors while course continues. Soon after trainees complete their MTNPs, copies of the instrument should be sent to their supervisors. A cover letter, signed by a respected program official, should explain that the evaluation will help the program to offer more job-related courses and, thus, help agencies maximize return on their training investment while increasing managerial proficiency. The due date for return of completed MTNPs should be shortly before trainees return to their jobs.

g. Acknowledge supervisor responses and explain mechanics of follow-up to them. Because of the common problem of getting supervisors to cooperate with such evaluations, every effort should be made to win the supervisor's continued interest. Supervisor respondents should receive a letter of acknowledgment and thanks that also details the follow-up's mechanics.

h. Remind nonresponding supervisors of the need for cooperation. If time permits, nonresponding supervisors should be reminded, by mail or telephone, of the potential value of cooperation. The reminder should occur appreciably before the scheduled due date. A second copy of the MTNP might also be sent to nonrespondents.

i. Review supervisors' MTNPs and collect incomplete data by telephone. Supervisors who return the MTNP do so under less duress than trainees, so their forms tend to be more complete. If short parts happen to be left blank, these should be completed by telephone. Forms that were apparently completed after trainees returned to their jobs should not be counted as pretraining data. This is because supervisory perceptions of trainee proficiency will be colored by training results and, thus, cannot validly reflect pretraining needs.

3. Collect Posttraining Proficiency Data. Because it is done by mail, with limited telephone contact, the follow-up data collection is extremely straightforward.

a. Distribute MTNPs to trainees and supervisors at the agreed-upon time. Forms should be mailed out at the two- to three-month date that respondents anticipate.

b. Stress the value of participation in the evaluation. Again, the potential values of participation should be pointed out. For trainees, these include self-knowledge, helping the program staff to improve their offerings, and seeing to it that future trainees experience optimal job-relevancy in the course. For supervisors, these include maximizing return on their training investment and further improving performance of trainees who are sent to the course in the future.

c. Give the option of obtaining a summary of results. Trainees and supervisors alike should be given the option of later getting a two-page summary of evaluation results. This helps increase return rates.

d. Remind nonrespondents of the need for cooperation. Unlike the pretraining data collection from supervisors, the follow-up is performed without the time pressure of the trainee's imminent return to the job. Hence, there should generally be time for sending nonrespondents a reminder letter on or shortly before the due date. A second copy of the MTNP might also be sent, especially to those who responded to the pretraining survey.

e. Review responses and collect incomplete information by telephone. Short parts of the MTNP may again be completed by telephone.

4. Analyze Data. The basic analyses of pre- and posttraining data employ the OPM computer package, but comparisons of pre- and posttraining data can be as simple or complex as available resources, programmer creativity, decision-maker interests, and the quality and quantity of data permit.

a. Determine training needs reported before training. For both groups, trainees and superiors, the OPM computer package produces group reports with seven tables showing (1) the number of managers needing training in the 14 activity areas and the training priority for each area; (2) knowledges and abilities in which 50 percent or more need training; (3) number of managers needing training in each of 61 knowledges; (4) number of managers needing training in each of 50 abilities; (5) levels of training needed in each activity area; (6) levels of training needed in each knowledge; (7) levels of training needed in each ability. It also generates training needs profiles for each individual manager, seen from two perspectives.

b. Determine training needs reported after training. The same analyses performed on pretraining data are peformed on follow-up data, producing two group reports reflecting trainee and supervisory perceptions.

c. Compare pre- and posttraining data. If the surveyed population is large enough and other conditions make it appropriate, a two-by-two analysis of variance should generally be performed on all key proficiency indicators. There is no hard and fast rule on the size population needed to justify doing fairly sophisticated statistical analyses; the decision as to how results will best be analyzed hinges on both the size of the universe to which results are to be generalized and on the magnitude of change that one could realistically expect from the course.

The most basic analyzes, which should always be done, involve the training priority ratings for the 14 activity areas. First, we should see if trainees and supervisors had compatible pretraining views of training needs. By calculating a Spearman's correlation or similar statistic, we can determine whether trainees and supervisors ranked training needs similarly before training. Second, we can see whether they still agreed after training. Then we can examine the pre- and posttraining data for each of the two groups. First, we can see if trainee rankings of training priorities have changed. Second, we can examine supervisory rankings. Then it may be useful to calculate a Pearson's Rho, to see if trainees and supervisors changed their perceptions of training needs in roughly similar ways.

To determine whether and how the course succeeded in reducing training needs, we may start by examining the training priority scores for the 14 activities (not their ranks). By visually scanning them, we can gauge whether and how much the magnitude of training priorities has dropped. Pre- and posttraining priority scores should be independently compared for trainees and supervisors. A fairly consistent decline in priorities generally signals the utility of further examination of results. A consistent drop in the number of knowledges and activities in which 50 percent or more of trainees are deficient similarly suggests the usefulness of more rigorous data analyses, if the size of the surveyed population and other conditions warrant.

In any further analyses--whether analysis of variance, t-tests, or other techniques are used--the nature of the course needs to be taken into account. This is important because many courses try to accomplish purposes other than immediate increases in proficiency. For example, a course might be equally directed toward changing trainee views of job requirements and toward increasing current capability levels. If this were not taken into account, the evaluation might overlook that (1) the course is succeeding quite well in getting trainees to redefine their job requirements upwards and (2) proficiency has also increased, if the lower pretraining job requirements are used as the basis for determining both pre- and posttraining proficiency.

d. Assess results in relation to other evaluation data. MTNP results should not be viewed in isolation but should be compared with the results of any other analyses. In some cases, it may be valuable to compare individual data collected through several means (e.g. MTNP, follow-up administration of final exam to test long-term retention, and utilization surveys).

5. Report Evaluation Results. Due to the complexities of the data generated by this approach, it is generally advisable to present evaluation results to program decision-makers orally, rather than in the traditional evaluation report.

a. Brief program decision-makers. The decision-maker briefing should first review the evaluation's purposes and data collection procedures, illustrate what data the group reports contain, and then outline the questions that guided data analyses. Then group data should be presented to answer those questions. Typically, there will be three questions of interest:

- How much agreement is there between trainees and supervisors in their perceptions of training needs? Are their priorities compatible? If the priorities change, do they change in similar ways?
- What indications do we have that training needs have been generally reduced through training?
- In what specific course areas is there most evidence that training needs have been ameliorated? In what areas is there most evidence of continuing need?

The briefing should be supported by handouts and graphic illustrations. It should also be open for questions. In other words, the presenter may find out that certain variables have been overlooked and that further analyses should be done. For example, the presenter could find out that job requirements require special attention, as we suggesed in an earlier illustration. As a result of the briefing, program decision-makers should at least have a sense for (1) whether their target population experiences communication problems about job requirements, (2) how well the course is generally reducing training needs, and (3) in what ways the course could be improved to be more job relevant. Ideally, they will also articulate the implications of evaluation results for further action in adjusting the course. Also as a result of the briefing, the presenter should know (1) whether further data analyses are needed and (2) what information should be included in the final summary of evaluation results.

b. Develop short summary of evaluation results. When further analyses have been completed and any additional briefings held (if warranted), the evaluation's results should be summarized in a short report. Typically, this report should be put in letter form, to maximize chances that it will be used and to deter evaluators from lengthy discourse. The letter should highlight the evaluation's purposes, data collection techniques, questions asked, results, and implications. Any new data generated after the last briefing should be included but without detracting from the letter's primary focus.

C. SIMPLIFIED COST-BENEFIT ANALYSIS

The Simplified Cost-Benefit Analysis (SCBA) Approach--which OPM calls Value Model II--builds directly on the PA approach and goes a step beyond it by adding a cost component. It addresses the central question: Is training a reasonable investment to make, based on its return to the organization in increased job proficiency? It can help answer several specific questions, such as:

- What is the economic worth of a position to the organization?
- What is the relative worth of each of 14 managerial activity areas to job proficiency?
- What are the costs to the organization of pretraining and posttraining deficiencies in the 14 activity areas?
- What is the value to the organization of improvements in proficiency brought on by training?

The concept underlying SCBA is that every managerial deficiency constitutes a cost to the organization. We can look at this as the organizational cost of doing nothing to bring a manager's capabilities up to a position's full requirements. For training to be a wise investment, it should reduce the costs to the organization of the manager's deficiencies in a greater amount than the costs of sending the manager away to training. So, over and above proficiency data, SCBA presumes the need for three other types of information: the costs of supporting the manager's position, the costs of sending the manager away to training, and the amortization period over which the organization can expect to benefit from training.

The preconditions for and cautions about the use of SCBA are primarily the same as for PA, on which it builds. We add only a few other points. This approach can be best used to determine if training's benefits outweigh its costs for programs that try to increase proficiency in ways that it is hard to attach a dollar value to. It should not be used when the dollar value of improved proficiency can be determined directly. The accessibility of decent cost data may also determine whether SCBA is appropriate. Determining the costs of a manager's position and of sending a manager to training can be complex. It is especially difficult to generate comparable cost data for an entire class, when trainees hail from several widely dispersed agencies of varied size and with different accounting systems. The amortization period, by which any cost savings would be multiplied, is typically all but impossible to estimate reliably unless careful follow-ups have been done previously. These cautions should be taken into account in deciding whether to try SCBA or to settle for the proficiency data alone. Like PA, SCBA requires no major changes in course orientation or structure.

Because this NEP study did not allow us to experiment with SCBA directly, we are not in a position to describe the evaluation process in a level of detail comparable to APFUA and PA. The description should provide sufficient detail, however, to decide whether the approach has enough appeal to explore it further. The SCBA overview contains five general steps. Supplementary resources for SCBA are the same as the ones listed for PA.

1. Determine Pretraining Proficiency Levels. Pretraining proficiency should be determined on the MTNP, as in PA. In intially selling the evaluation to trainees, the reasons for needing cost data should also be explained, unless cost data can be obtained directly from supervisors.

2. Determine Costs to the Organization of Pretraining Deficiencies. There are three steps to determining the costs of pretraining deficiencies: finding the economic worth of trainee positions, apportioning economic worth over the 14 activity areas according to their importance, and then calculating the actual costs of deficiencies.

a. Find the positions' annual economic worth to trainee organizations. The MTNP can be used to collect information on trainee annual salary. For a fuller picture of annual economic worth, other nonsalary indicators should be included in the estimate of worth: fringe benefits, overhead to support the position, salaries of subordinates, and so forth. Because there are several distinct methods for determining nonsalary worth, the critical consideration is maintenance of a consistent approach.

b. Apportion economic worth of positions over the fourteen activity areas. The MTNP collects importance weight on a scale from 0 to 50 for the 14 activity areas along with ratings of the required level and current level. To determine the annual economic worth for each distinct activity area, the importance weight for each activity as a percentage of the total weights for all 14 activities should be multiplied by the economic worth of trainee positions. The dollar figure represents the value to the organization of each activity's performance at the required proficiency level.

c. Calculate the costs to the organization of trainee pretraining deficiencies. The cost of a deficiency is the economic value of a benefit that the organization fails to receive. The cost of a trainee's pretraining deficiency in a given activity area is calculated by multiplying the trainee's relative deficiency for the activity by its economic worth. Relative proficiency is the ratio of current level to required level, up to a maximum of 1. The sum of the costs for all deficiencies represents the cost in lost benefits of doing nothing to correct them.

3. Determine Posttraining Proficiency Levels. Just as before training, the MTNP is used to have trainees and their supervisors rate the 14 activity areas in terms of importance, current level, and required level. If posttraining proficiency is no greater than the pretraining level, then the training has had no apparent benefit. The analysis should then skip to determining the costs of training to the organization and training's net loss. If, however, proficiency levels have generally improved, the value of this improvement should be calculated step by step.

4. Determine Training's Value. Determining training's value to the organization involves five steps: calculating training's annual gross value, setting an amortization period, factoring in the amortization period, determining training's costs to the organization, and calculating training's net value.

a. Calculate training's annual gross value. The right way to calculate training's annual gross value in a given activity area depends on whether the organization can benefit from training above the required level. If it is known that an organization cannot benefit from proficiency in an activity area above the required level (a relative proficiency ratio of greater than 1 to 1), then training's annual gross value in the activity area is calculated this way: the percentage increase in relative proficiency (up to a 1 to 1 ratio) is multiplied by the cost of the original deficiency. If it is known, however, that the organization can benefit from proficiency greater than the required level, the activity area's annual gross value is calculated by multiplying the percentage increase in relative proficiency (even above a 1 to 1 ratio) by the costs of the original deficiency. In most cases, this will not be known, unless the program has done extensive follow-up research on its effects, so either calculation procedure may be chosen and used consistently. The sum of the annual gross values for the 14 activity areas equals training's total annual gross value.

b. Determine the amortization period. Determining the amortization period for training (the number of years the organization will benefit from training) requires the prediction of several uncertain events. These include the trainee's departure from the organization, transfer to another position in the organization that does not call for what was learned in training, forgetting what was learned, obsolescence of the learning, and so on. Because typically no ready formula exists for determining these events, the only simple solution is to select alternate amortization periods.

c. Factor in the amortization period to determine total gross value. The amortization period(s) should be multiplied by training's total annual gross value to obtain training's total gross value. The present value of the dollar investment can be taken into account to ensure that the inflation factor is not overlooked.

d. Determine the costs of training to the organization. The costs of training to the organization fall into one or more of four categories: trainee costs of salary and benefits, travel, per diem, tuition, and materials and supplies; instructor costs of salary and benefits, overhead, travel, and per diem; facilities and equipment costs; and development costs of salary and benefits, overhead, travel, and per diem for course developers, and costs of contracts and development materials. Because no organization is likely to have costs in all four categories, relevant costs should be identified and totaled.

e. Calculate training's net value. Training's net value to the organization is calculated by subtracting the costs of training from the total gross value of training.

5. Report Evaluation Results. It is preferabe to follow the same course in reporting results as recommended for PA: presentation of results to program decision-makers orally. There is a much wider direct audience for the SCBA results, which includes the agencies that send trainees to the program. For these user groups, the critical question is not "how can we improve our course?" but "how can we get the greatest return on our investment?" These user groups cannot generally be reached by an oral presentation, unless the methodology has been modified by selecting a study population from a single agency that has attended several different courses. As in the typical evaluation, user groups should be contacted directly through a brief written report. This report may be in the form of a letter, but often this will not be appropriate because extensive supporting data will then have to be appended as tables. We recommend a brief, conventional evaluation report in communicating with user agencies.

The interpretation of SCBA results can be straightforward, if total net value greater than 0 is considered necessary and sufficient evidence of cost-effectiveness. If opportunity costs are taken into account--these include the costs of other investments not made, such as other types of training--then the total gross value needed to provide convincing evidence of cost-effectiveness may be set higher. In contrast, when the analysis shows training has had little or no benefit (posttraining relative proficiency is no greater than it was before training), then there is more room for interpretation: the course itself does not provide instruction that fits organizational needs, or the wrong trainees are being sent to training, or the user agenicies do not permit trainees to display greater capability, and so forth. At the minimum, a negative finding throws a red flag, calling for closer scrunity of the value of continued investment in a program.

* * * *

This chapter has outlined three approaches for examining the central policy- and funding-related questions faced by police management training programs. In Part 3 of this report, we review the evaluation literature and consider ways to research specific national policy questions.

PART 3: RESEARCH INDICATIONS AND GENERAL FINDINGS

From the first two parts of this report, one can understand how practices in police management training program development compare with the industrial model. One can appreciate the external factors that impede systematic program development. One is familiar with the range of models by which these programs operate, can assess a given program in terms of the seven evaluability criteria and identify potential ways to make it more evaluable, and also knows several ways to measure training's effects on later job behavior.

This final part of the report starts by taking a more traditional evaluation approach. It asks three questions:

1. How have police management training programs been evaluated, and what conclusions about program effectiveness may be drawn from the available evaluation literature?

2. What national policy questions about police management training go unanswered, what research approaches will allow us to answer them?

3. What conclusions may be drawn from the study, what actions should be taken on the basis of these findings?

Chapter Seven addresses Question 1 by demonstrating the types of conclusions that may be drawn from the limited evaluation literature. Because programs differ so much in their purposes, it is difficult to draw conclusions about police management training programs in general. Rather, most evaluation findings have relevance that is limited to the type of program, the level of participants, and, to an extent, the specific evaluated program. We examine three different evaluation categories: follow-up utilization surveys, incumbent and superior performance ratings, and attitudinal self-assessments. From this review, we conclude how strong the evidence is that trainee learning is relevant to their needs, fits the organizational context to which trainees return, and has measurable payoff in greater individual and organizational performance.

Chapter Eight explores Question 2 by showing how several unanswered policy questions in this area could be studied. These range from the cost-effectiveness of residential vs. non-residential programs to the relative effectiveness of experimental and non-experimental programs. For each of the questions considered in depth, it outlines a research design and discusses the policy implications of different potential findings. In addition, it lists several other policy questions that warrant further study.

Chapter Nine concludes the report by summarizing the major study findings and conclusions. It also recommends to oversight agencies, program managers, program operators, police agency executives, program evaluators, and researchers several means for learning more about the operations and effectiveness of police management training programs while also making them more evaluable and, hence, more manageable.

These three chapters draw upon little of the study's empirical data, with the exception of Chapter Nine. Chapter Seven is based primarily upon published evaluation literature, although it also taps evaluation practices observed on site or reported in the national mail survey. Chapter Eight draws upon the views of all interested parties--especially oversight agencies and program managers--to identify the unanswered questions that are important and warrant study. The research designs outlined in this chapter do not build directly on any data collected for this study, however. Chapter Nine, because it summarizes the study's conclusions and recommendations, builds on all the study's data sources and on prior chapters of the report.

Seven: Evidence Of Program Effectiveness

What evidence is there that police management training programs are effective? Because of the many obstacles to evaluating these programs and the consequent reluctance of programs to invest in extensive evaluations, the evaluation evidence lets us draw few conclusions with confidence about how such programs are working. The literature offers scant evidence showing that the knowledge and skills trainees learn are relevant to their personal needs, fit the organizational context from which they came and to which they return, and have measurable payoff in greater individual proficiency and organizational effectiveness.

The bulk of the evaluation literature consists of reaction surveys and final course examinations. These are better regarded as assessments rather than evaluations, if we define evaluation as a systematic effort to draw inferences from several sources of data in order to determine a program's value. So understood, an evaluation will minimally include some measure of in-program outcomes and of transfer to the job situation. We do not consider reaction surveys and knowledge exams alone to be forms of evaluation evidence because they are useful only in making differentiations within (not across) programs. They become evaluation evidence when they are incorporated into a larger evaluation that gathers several types of evidence (which is usually not the case). The plentiful supply of such assessments tells us convincingly that:

- Trainees leave virtually all programs, regardless of variations in quality, satisfied that their time has not been wasted
- Trainees leave programs with as much or more knowledge of the rudimentary concepts dealt with during the program session than they carried into it

This chapter's purpose is to demonstrate the other types of conclusions that may be drawn from the evaluation literature. We do not intend to dissect critically these few ambitious efforts to document training's value, but we do point out certain weaknesses and state the level of confidence that one may place in an evaluation's results. To the extent that we appear critical, our comments are not directed toward the evaluated programs themselves, which include some of the most renowned in the country. We discuss three categories of evaluations: follow-up utilization surveys, incumbent and superior ratings of performance, and attitudinal self-assessments using standardized personality inventories.

The results of this study's efforts to pretest certain evaluation approaches, discussed in Chapter Six, are not included in this chapter. Had we included them, they would have fallen under follow-up utilization surveys and incumbent and superior performance ratings.

A. FOLLOW-UP UTILIZATION SURVEYS

Most follow-up utilization surveys fall at the weak end of this spectrum because they do not focus on changes in job behavior brought about by training but, instead, concern more realistic trainee estimation of utilization's likelihood. Most, therefore, are more appropriately considered follow-up reaction surveys or "rebound evaluations" rather than true utilization surveys. These follow-up reaction surveys can still be used to put immediate post-program reactions in perspective and, in comparison with them, can provide useful information about how well program philosophy and contents relate to the organizational environment. Unfortunately, few programs systematically compare short- and long-term trainee reactions. The inherent limitation of even such follow-up reaction surveys is that they do not deal with the key utilization question: What do you do now that is different from what you did before? Stated somewhat differently: Have you been able to use what you learned to adopt new behaviors? Two major evaluations focus more directly on utilization: Rush's study of California's acclaimed police middle management program (PMMP) and Hettinger's evaluations of the LEAA-funded Executive Training Program in Advanced Criminal Justice Practices (ETP).

1. California's Police Middle Management Program. Rush's study examined the question: Does attendance of police middle managers at middle management training make any differences in the later management and organization of their agencies?19/ Rush hypothesized that, due to organizational constraints, training would have little impact on the police organization and that the police organization, characterized by a semi-military command structure and a rigid chain of command, would resist change and make it difficult, if not impossible, to implement course concepts. The assumption behind this hypothesis is that organizational change is a "herculean task," in which police agencies face the central problems of generating a positive organizational climate and initiating the process of change.

To answer this question, Rush obtained the personal perceptions of 436 graduates from 41 sessions of PMMP conducted from 1969 to 1974 at three of the eight institutions certified to offer the course. The survey respondents, who represented 125 municipal agencies, completed a forced choice, close-end questionnaire that covered 220 variables. Survey questions dealt with individual background, perception of the middle manager's role, organizational influence, individual ability to initiate change, and the impact of the PMMP course on the trainee and the trainee's organization. Data were analyzed for frequency on key background variables (rank, age, education, and so on) and t-tests were calculated based on the means.

The results of the study convincingly supported the hypothesis that middle management training, for certain correctable reasons, does not have much impact on the organization of police agencies. The vast majority of respondents

19/ G. E. Rush, An Evaluation of Police Middle Management Training in California (Unpublished doctoral dissertation, Claremont Graduate School, 1975).

indicated that their organization either had not changed at all or had changed only slightly as a result of PMMP attendance. The study contained several conclusions, but these were key:

- Graduates are unable to implement concepts from training and alter their organizations, apparently because superiors are not familiar with the proposed changes or because top administrators are reluctant to accept constructive advice. Middle management training, however, does not address the predicament of the middle manager, whose position is not one of "motion maker," but lies in the "murky middle of the organizational milieu" as the "organizational hemostat maintaining equilibrium between management and those managed." Management concepts are beneficial only if "set in motion," and, to this end, middle managers are in a "very precarious position."

- Graduates who have a positive attitude toward their organizations place greater value on the PMMP course. Specifically, those who said their departments had an interest in helping them obtain promotions, were receptive to their suggestions, held group problem-solving meetings, and assisted in decision-making skill development placed greater value on the course. Those who are more comfortable with their organizational environment are more receptive to organizationally oriented training. This relationship between trainee attitudes and perceptions of course value means that ultimate course "success" in influencing other officers through trainees hinges on an uncontrolled, external variable.

From these conclusions, Rush made several major recommendations about program delivery and evaluation. First, the evidence that middle managers do not see that PMMP attendance has influenced their organizations suggests that all police managers, at all management levels, should take part. In this way, all would be exposed to the same concepts, the "mysteries" of management eliminated, and an environment more supportive of change created. Second, the negative attitudes of trainees toward their organizations and toward organizationally oriented training programs can be influenced through training if evaluators take them into account in assessing the outcomes of training and if trainers deal with them directly in training activities.

To appreciate the conclusions of Rush's study, it is important to note that the California PMMP, as it existed at that time, was a mixed model, combining three basic models from Chapter Four: systematized policing, state-of-the-art, and non-experiential participative. It is not clear how much the program was in flux from 1969 to 1974, when the study population attended PMMP, but since then, it has shifted dramatically into an experiential participative model, though still mixed. Rush's evaluation later influenced not only the California program but also LEAA's decision to direct its major criminal justice training program, the Executive Training Program, toward those with clear authority and decision-making power to implement practices proved effective.

2. LEAA's Executive Training Program. Hettinger's evaluations for ETP's first two years (1977 and 1978) have tried to determine whether trainees with direct responsibility for implementing expected changes were succeeding in their efforts to implement personal end-of-course action plans.20/ To understand the context of these evaluations, we note that ETP's goal of disseminating or transferring specific innovative yet time-tested techniques in focused problem areas differentiates it from the broader, less defined purposes of a program like California's PMMP, which has no pretensions of giving participants a roadmap to management. In this sense, ETP can be considered more operationally oriented than programs emphasizing management concepts. We also note that ETP is directed toward a wide criminal justice audience, with only selected topics offered to police managers in particular.

To determine the extent and areas of implementation, Hettinger drew appropriate samples for each program (or workshop topic). Participants were followed up to determine: how much of the end-of-course action plan had been implemented, the specific areas in which changes were made, steps taken as a result of the work shop, and consequent effects on the agency. The action plan is different from that described in Chapter Six in that trainees have no control over the overall design of the plan. The action plan consists of a checklist on which trainees mark what they will try on the job. In the followup, respondents indicate the extent to which they implemented the plan on a scale from 0 through 100, in intervals of 25. They express changes by checking forced-choice alternatives. The ETP program evaluations uniformly indicate that nearly all participants succeeded in implementing at least a portion of their action plans, with varying percentages trying to change components or take steps.

There are problems in interpreting ETP results, aside from the lack of checks on the validity of responses. First, the extent of implementation that LEAA considers necessary to indicate success is not clear. The ETP evaluator seems to view "Some," or 25 percent, implementation as success, although this is their lowest category out of five, other than "Not at all." In contrast, Rush viewed implementation of "one or two concepts" as an indicator of the lack of organizational impact, although this was also the lowest category out of four, aside from "none." Second, the distinct orientation of ETP, directed toward implementation by top managers of practices that are spelled out in an unambiguous manner and packaged in take-home documents, suggests that findings about its success may have limited generalizability. Later annual ETP evaluations by Hettinger and an ongoing ETP evaluation by Rae of the Institute for Social Analysis may help clarify these questions.

B. INCUMBENT AND SUPERIOR PERFORMANCE RATINGS

The best known application of incumbent and superior performance ratings took place in the LEAA-funded National Manpower Survey of the Criminal Justice

---

20/ B. Hettinger, Summative Evaluation of Cycle I, Regional Training Workshops (Washington, D.C.: University Research Corporation, December, 1977); Summative Evaluation Report, Regional Training Workshops, Fiscal Year 1978 (Washington, D.C.: University Research Corporation, September, 1978).

System (NMS).21/ NMS examined the "sufficiency" of several types of police training, up to the supervisory level. It used three primary methods to determine whether existing supervisory training programs were sufficient. The duration and content of training were compared with national standards. Incumbents were asked whether they had learned various job-related tasks primarily through formal training or primarily through on-the-job experience. Police executives and knowledgeable incumbents were asked about the extent to which trained newly assigned personnel have requisite expertise. We will be concerned only with the latter two approaches.

To determine whether incumbents learned job-related tasks through training or job experience, NMS asked a nonrepresentative sample of 165 patrol supervisors and 96 detective supervisors from 31 agencies whether they learned to perform the tasks that task analyses showed were involved in their jobs from training or from job experience. A majority of interviewees said they learned all the tasks specific to their jobs through on-the-job experience. In two particular areas related to management, the majority said they had been inadequately prepared. Incumbents felt training inadequately prepared them to perform personnel and administrative actions such as evaluating job performance, taking disciplinary action, recommending awards, approving promotions, approving directions, and counseling subordinate personnel. They also felt ill-prepared to conceive, plan, and recommend improvements, innovations, and changes in department policies, objectives, and procedures for coping with crime and providing public service.

To determine whether executives and knowledgeable incumbents perceived trained newly appointed personnel as having appropriate expertise levels, NMS asked 90 respondents to review a list of skills and knowledge topics relevant to the two supervisory jobs and to rate the level of knowledge or skill required and the level attained by newly promoted incumbents on a five-point scale. Interpreting a one-and-a-half rating points difference between required and attained levels to indicate a considerable deficiency, new appointees were found to be deficient in several important areas, some related to supervisory and administrative responsibilities.

NMS appropriately interpreted these findings with caution, as a limited indication that current training is not sufficient to bring newly appointed supervisors to a required level and that more and/or better training is needed. NMS qualified its conclusions by noting that the small study sample cannot be shown to be representative and that the study cannot offer data on the pretraining performance of new appointees. NMS also admitted the lack of evidence that experienced incumbents who took part in rating appointees meet required levels and the lack of clarity as to training's capacity to raise expertise levels significantly.

21/ National Planning Association, National Manpower Survey of the Criminal Justice System -- Volume 2: Law Enforcement (Washington, D.C.: National Planning Association, October, 1976), pp. 222-226.

C. ATTITUDINAL SELF-ASSESSMENTS

Two originally LEAA-funded programs--the New England Institute of Law Enforcement Management (NEILEM) at Babson College and POLEX at the Pennsylvania State University--were evaluated in terms of long-term effects of program attendance on deeply embedded attitudes. Seen side-by-side, the two evaluations look remarkably different: the one (NEILEM) writing off the prospect of attitude change as implausible and the other (POLEX) underscoring the implications of expanded training opportunities for the promotion of executive leadership qualities. What makes their divergence all the more noteworthy is that the programs bore certain resemblances at the time they were evaluated. Both represented a mix of three basic models: systematized policing, state-of-the-art, and non-experiential participative. The study populations for both included graduates from the programs' first years, when program practices and orientations were probably less stable than in later years. Due to reasonable program similarities, differences in conclusions seem likely to be a function of method. A third evaluation of a POLEX spinoff program is compatible with the POLEX results.

1. New England Institute of Law Enforcement Management. Hornaday and Kaiser tried to test the hypothesis that "command training officers' managerial behavior can be changed by a training course designed to increase their acceptance of responsibilities, their understanding of the nature of human motivation, and their skills in techniques for managing people."22/ This evaluation does not explicitly set forth hypotheses about effects of training on personality, but it seems to dismiss personality change as implausible, stating, "The assumption was that the program could change and improve behavior but not basic personality.... [because] there is no reason to expect that a three-week course would or could produce a personality change. In fact, it is desirable that the course not change the personality of participants." On its face, this position seems incompatible with the evaluation proposal's plan to examine the effects of training on personal accomplishment, judgment by immediate superior, motivation (drive), character and personality, and leadership style. Although it is evident that the evaluation's measures of success in part changed (e.g., judgment by immediate superior was not included), it is not clear whether hypotheses actually changed or were simply never articulated. In either case, the evaluation's follow-up reaction survey plays a more central role in demonstrating alleged changes in behavior than one could anticipate from the evaluation proposal.

To determine the effects of training on personality, Hornaday and Kaiser compared results on five tests administered to 881 program graduates from 1967 to 1971 and, in the absence of pretraining measures, to those of a control group of 295 that was matched for state and agency represented, size of community, and rank. All respondents were experienced officers from New England states, with the rank of corporal through chief, and most were from municipal agencies.

22/ J. Hornaday and R. Kaiser, The Effectiveness of a Managerial Training Program for Command Law Enforcement Officers (Wellesley, Massachusetts: New England Institute of Law Enforcement Management, 1972).

Respondents took four personality inventories and a personal interest test: Gordon's Survey of Interpersonal Values, Gordon's Survey of Personal Values, Fleishman's Leadership Opinion Questionnaire, Kuder's Occupational Interest Survey, and Wonderlic's The Personnel Test. We will not comment on the different personality constructs these tests measure and their validity as measures except to note that the constructs measured are varied (include leadership motivation) and the tests widely used. Hornaday and Kaiser performed t tests on their results, differentiating between control and experimental groups and by rank and municipal/state police. Except for a few scattered significant findings, these comparisons revealed only nonsignificant differences. We tend to concur with the authors that personality change as measured by these tests and procedures would have been a surprising finding. We concur because of the program's short duration and because of the the lack of correspondence between the program's objectives and the personality measures used.

2. POLEX. The well-documented evaluation by Newman, Price, and Horner of POLEX's effects on leadership-related personality aspects differs from the Hornaday and Kaiser evaluation of NEILEM in several key respects.23/ In terms of evaluation design, the POLEX evaluations:

- Explicitly hypothesized a short-term positive effect of program participation on personality traits related to leadership and the long-term stabilization of these short-term positive changes
- Intentionally selected as its personality measure an inventory that included scales already externally validated for their relationship to leadership and that were, consequently, plausibly related to POLEX goals
- Obtained measures of personality from the study participants not only after a long posttraining duration but also immediately before and after training

We speculate that the POLEX evaluation, for which data collection began in the fall of 1971 with the commencement of POLEX, may have indirectly profited from the initial conceptualization and evident shortcomings of the NEILEM evaluation, funded a few months earlier. Whether or not this is so, the POLEX evaluation somehow overcame the principal shortcomings of the NEILEM evaluation.

The POLEX researchers collected data from 127 police middle managers and executives in 1971-1972 and from at least another hundred over the next year. Participants were tested on Grygier's Dynamic Personality Inventory (DPI), a

23/ Readily accessible articles have appeared in two journals. See C. Newman, B. Price, and J. Horner, "Police Executive In-Service Training and Its Effect on Selected Personality Traits" (two parts), Police Law Quarterly, Spring, 1973, pp. 14-27, and Summer, 1973, pp. 42-47; also, C. Newman and B. Price, "Police Executive Development: An Educational Program at the Pennsylvania State University," Police Chief, 41 (April, 1974), pp. 74-77.

384-item questionnaire with 34 subscales, of which seven are empirically related to leadership. Testing occurred immediately before and after participation in the four-week institute. In addition, some early classes were retested a year after graduation. Analyses were performed on four of the seven subscales that were related to leadership and six others that were stereotypically related positively or negatively to police officers. (The six were included in the analysis because of the researchers' parallel interest in a question unrelated to our proposes: What are the differences between personality characteristics of police executives and the general population?) Data were analyzed using t tests and univariate analyses of variance. The results showed that, on three of the four traits related to leadership, the responses of participants immediately after training were significantly different from pretraining responses. Limited follow-up data supported the hypothesis that, after participants return to their jobs, the posttraining leadership improvements stabilize.

3. POSIT. A closely related study, conducted by Price and Adelberg to assess the effects of a POLEX spinoff program for supervisors (called POSIT), substantiates the POLEX findings. It also extends them to a lower management level, applies more appropriate analytic techniques, and sheds light on the nature of the change the takes place.24/ Price and Adelberg tested POSIT participants immediately pre- and post training on a slightly abbreviated version of the DPI, called the Likes and Interests Test (LIT). The evaluation examined the effects of training only on the seven LIT traits related to leadership. These are more related to program goals and, hence, a more appropriate measure than some of the traits used in the POLEX evaluation. In addition, the evaluation recognized that, because the traits in the LIT, like the DPI from which it was adapted, are interrelated, the use of individual tests of significance as in the POLEX evaluation would be inappropriate. Instead, the POSIT evaluation examined the seven traits related to leadership using Hotelling's T Square Test, the multivariate analog of the t-test.

Analyses showed an overall improvement in the seven leadership-related traits, but Price and Adelberg were not satisfied with these results. To interpret these changes more clearly, they then performed a factor analysis from which two factors emerged. These factors they labeled leadership initiative and leadership maturity. Leadership initiative, on which creativity, emotional independence, drive for achievement, and verbal aggression loaded highly, accounted for 74 percent of the variance. Leadership maturity, on which social activities, persistence, and initiative loaded highly accounted for only 25 percent of the variance. Further analyses corroborated that there were significant changes in the first cluster of traits but not in the second. From the results, Price and Adelberg concluded that it would have been implausibele to expect impact on the second cluster, leadership maturity, since leadership style is slow to change; in contrast, in the first cluster, leadership initiative, is more subject to change based on external stimulation. They further concluded from these results

24/ B. Price and S. Adelberg, "An Evaluation of Police Supervisory Training Using a Multivariate Assessment of Attitude Change," Journal of Police Science and Administration, 5 (March, 1977), pp. 69-73.

that POSIT succeeds in "motivating its participants to more fully assume leadership responsibilities and increases their confidence and willingness to work effectively in the supervisory role." Price and Adelberg thus provide the basis for setting more plausible training expectations and for better interpreting resultant personality change.

* * * *

This review shows that the evaluation literature offers scant evidence that the knowledge and skill trainees learn are relevant to their personal needs, fit the organizational context from which they come and to which they return, and have measurable payoff in greater individual proficiency and organizational effectiveness. Moreover, it shows that few systematic evaluation strategies have been implemented to answer these questions. In the next chapter, we look at several research designs for answering questions with major policy implications.

Eight: Research Approaches To National Policy Questions

What research approaches offer answers to the important national policy questions in this area? There are a dozen or more such questions, and the future structure, practices, and orientation of police management training hinge on whether and how they are answered. With that in mind, these questions should be the focal point of future research. This chapter's purpose is to summarize general research approaches to several of them in a way that also suggests additional approaches. Research approaches to the following six policy questions are considered in depth:

- Do residential programs have greater effects on knowledge and proficiency than nonresidential programs?
- Are residential programs cost-effective in comparison with nonresidential programs?
- What effect does a network of program graduates have on the management practices, career development, and interdepartmental relations of network members?
- Do programs that use experiential teaching methods have more influence on later job behavior than programs that cover similar material but rely exclusively on traditional lecture-and-discussion methods?
- Do graduates of major national programs experience more rapid career development and achieve higher terminal positions than comparable nongraduates?
- Does the pretraining cooperative development by trainee and supervisor of personalized learning objectives result in higher subsequent implementation levels than traditional trainee preparation?

Each research approach summary starts with a brief discussion of the problems and of the policy implications that would flow from different research outcomes. We then outline the independent and dependent variables under study, the treatment or intervention to which the experimental group will be exposed, the criteria for selection of programs and/or trainees, and the measures to be taken of the dependent variables. A few of the research approaches draw on Chapter Six's three evaluation approaches to measure dependent variables.

A. EFFECTIVENESS OF RESIDENTIAL VS. NONRESIDENTIAL PROGRAMS

Residential programs claim to be more effective in shaping trainee behavior. They base this claim on the extension of the learning process beyond the classroom into off-hours informal interaction among trainees, aided by the extra

opportunities for formal and structured off-hours activity. The differential effects of the residential environment on training's effectiveness, however, are not known. Do residential programs have greater effect on knowledge and proficiency than nonresidential programs?

If residential programs are more effective than nonresidential programs, trainees can be expected to learn more during the program, retain more of what they learned after they return to their jobs, and display greater changes in job proficiency. If residential programs are shown to be more effective than nonresidential programs, the current movement away from residential programs in favor of state and local commuter programs will have to be reassessed. Program managers will obtain a justification to intensify the residential aspects of their programs. Police executives will have a rationale for projecting the potential value to be derived from a program prior to sending trainees. LEAA and other funding agencies will have evidence that prior support of such programs has been reasonable and should be maintained and expanded, where feasible.

1. Variables. The independent variable is participation/nonparticipation in a residential program. The dependent variables are knowledge immediately at the end of the program, retention of this knowledge, and job proficiency.

2. Treatment. Residential management training consists of the following activities in a residential setting: in-class lectures and discussions on significant management topics and problems, extended in-class trainee formal interaction, off-time trainee group interaction on assigned projects and classwork, informal off-time trainee social interaction, and extensive staff control and coordination of formal and informal trainee interactions.

3. Selection. The programs to be compared should be operated by the same organization, of similar duration, using the same instructors to convey similar material to trainees from similar departments, who are also comparable in terms of other factors. These potentially include experience, length of service, prior training, educational achievement, and time in grade. Selection of programs for comparison would not have to take into account the actual trainee characteristics. If feasible, trainees should be randomly assigned to the residential and nonresidential conditions instead of statistically equating the two groups. Because this will probably not be feasible, trainee characteristics may have to be statistically equated.

4. Comparisons. Within- and between-group comparisons should be made (pre- and posttraining; residential/nonresidential program).

5. Measurement. Learning should be measured on a pre-/posttraining administration of the course final exam. Retention should be measured by a follow-up administration of the final exam three months after return to the job. Proficiency should be measured by a pretraining and three-month posttraining follow-up administration of the Training Needs Profile (TNP).

B. COST-EFFECTIVENESS OF RESIDENTIAL VS. NONRESIDENTIAL PROGRAMS

Residential programs may be more effective than nonresidential programs in some respects, but they are also more costly. Costs of operating residential programs are higher, and supporting departments face greater expenses for tuition, per diem, travel, and callback difficulties. It is important for funding agencies, police departments, and program managers to establish whether residential programs are cost-effective, especially in comparison with similar nonresidential programs.

If residential programs are cost-effective in comparison with similar nonresidential programs, then the value to departments of resultant trainee proficiency improvements should offset the greater costs. Showing that residential programs are cost-effective would establish that departments were getting a fair return for their investment. This would promote increased utilization of residential programs, provide residential program managers with arguments for strengthening program stability, and help reverse the trend away from residential and toward state and local commuter programs. It would also provide funding agencies like LEAA with justification for continuation or resumption of support to residential programs. In contrast, negative results would reinforce a de-emphasis on residential programs and signal the need for sponsoring agencies to reassess their support of residential activities.

1. Variables. The independent variable is participation/nonparticipation in a residential program. The dependent variable is job proficiency seen in relation to the costs of course attendance and the economic worth of trainee positions to their departments.

2. Treatment. Residential management training consists of the following activities in a residential setting: in-class lectures and discussions on significant management topics and problems, extended in-class trainee formal interaction, off-time trainee group interaction on assigned projects and classwork, informal off-time trainee social interaction, and extensive staff control and coordination of formal and informal trainee interactions.

3. Selection. The programs to be compared should be operated by the same organization, of similar duration, using the same instructors to convey similar material to trainees from similar departments, who are also comparable in terms of other factors. These potentially include experience, length of service, prior training, educational achievement, and time in grade. Selection of programs for comparison would not have to take into account the actual trainee characteristics. If feasible, trainees should be randomly assigned to the residential and nonresidential conditions instead of statistically equating the two groups. Because this will probably not be feasible, trainee characteristics may have to be statistically equated.

4. Comparisons. Within- and between-group comparisons should be made (pre- and posttraining; residential/nonresidential program).

5. Measurement. Proficiency should be measured by a pretraining and three-month posttraining administration of the TNP. In accord with the Proficiency Analysis Approach described in Chapter Six, both trainees and their

supervisors can be asked to rate the required trainee skills and knowledge as well as actual performance levels. The economic worth of a position can be measured in accord with the principles of Simiplified Cost-Benefit Analysis in Chapter Six. Economic worth takes in the salary and nonsalary costs needed to support a trainee's position. The costs of a course to the sending agency involve tuition, travel, per diem, the cost of replacements, and other identifiable expenditures. The question of cost-benefit can be determined by examining the increased proficiency of residential and nonresidential trainees comparatively, factoring in costs to departments and the trainee's greater ability to perform at full economic worth, and then prorating these over the length of time for which the course's results are presumed valid. In effect, this study involves Simiplified Cost-Benefit Analysis for experimental (residential) and control (nonresidential) groups.

C. EFFECTS OF A NETWORK ON JOB BEHAVIOR

Many programs emphasize the need to prolong the training process beyond the point when trainees return to their departments. To extend their influence and maximize impact on job behavior, they devote considerable attention to developing and nourishing networks of program graduates. Some programs define network activities more clearly than in-program activities. Similarly, graduates often claim that network activities and the bonds established with other officers are more beneficial than the program's formal classroom instruction. Little is known, however, about the actual effectiveness of networks in the improvement of police management practice. What effects do networks have?

Based upon the claims of programs stressing a network of graduates, a network should have several effects on police management practice, including improved individual problem-solving capabilities, better interdepartmental relations, and the implementation of innovative management practices. If network programs can demonstrate these effects, they will have shown how to prolong the training intervention and to achieve results much greater than would be expected from in-program activities alone. This would provide the rationale for restructuring other programs in which a network might flourish. It would also suggest the need for policymakers and funding agencies to reassess network activities and results to ensure that they square with intended activities. In contrast, if network programs do not demonstrate these effects, program staff might feel obliged to shift their emphases back to activities and outcomes over which they exert more control. Similarly, a negative research outcome would force LEAA and other funding agencies to scrutinize network programs more closely and to ensure that proper emphasis is placed on controllable in-program activities.

1. Variables. The independent variable is participation/nonparticipation in a program emphasizing the development and maintenance of a network of graduates. Dependent variables include: management problem-solving style, the frequency and quality of interdepartmental relations, and implementation of innovative practices.

2. Treatment. Network programs usually include intensive informal trainee interaction and socializing, staff efforts to structure and control the training environment, and staff emphasis on the prospects for prolonging program activities through a network. Network programs are typically residential and of

relatively long duration, but neither of these is an essential characteristic. Post-program network activities include continuing social contacts among graduates, graduate participation in formal reunions and newsletters, and assistance to graduates by program staff and other graduates.

3. Selection. The program(s) selected should have an easily identifiable and ostensibly well-established network. Two groups of trainees should be selected: one or more classes of graduates from the network program and a comparable group of police managers who applied to and were found eligible for the program but who were not accepted or, for other reasons, did not attend. Ideally, qualified trainees would be randomly accepted/rejected and assigned to network/non-network groups. Because this will probably not be feasible, non-network subjects should be drawn from lists of eligibles. Some statistical equation of groups might be required.

4. Comparisons. Within- and between-group comparisons should be made (pre- and posttraining; network graduates and nongraduates).

5. Measurement. Network graduates should show changes in management problem-solving style that involve reduced insularity and increased assistance from officers outside the agency. This can be measured by content analysis of three-month Action Plan follow-ups and surveys of graduates and nongraduates about assistance obtained from other departments in problem solving. Improved informal interdepartmental relations may be measured through surveys of departmental staff and formal improvements through content analysis of departmental records. Implementation of innovative practices may be measured by three-month follow-ups on Action Plan results.

D. EFFECTS OF ANALYTIC VS. EXPERIENTIAL PROGRAMS

Managerial development can require changes in individual attitudes and behaviors far beyond new formal knowledge. Police management training programs have generally focused narrowly on formal presentation of substantive knowledge. Because of this stress on passive learning techniques, they have neglected direct methods for providing trainees with experiences in different management systems and for assimilating these experiences so as to change trainee attitudes and behaviors. Experiential programs build on the principle that "an experience is worth a thousand pictures" in attempts to improve trainee learning and change attitudes and behaviors. Thus, they use techniques that actively involve the trainee in the learning process. Little is known, however, about the differential effectiveness of passive and active managerial training methods. Are experiential programs really more effective?

If they are more effective, trainees can be expected to learn more during the program, retain more of what they learn after they return to their jobs, and display more behavioral and attitudinal change in their jobs. Positive research outcomes would underscore the need to pay more attention to contemporary learning theory in the design of management training programs. It would further demonstrate that programs can become more effective within existing resources by the incorporation of experiential exercises into lecture blocks and that this will make program results more measurable and, hence, more evaluable. It would also suggest that funding agencies require greater emphasis on experiential methods

by funded programs. In contrast, negative research outcomes would reaffirm the acceptability of most current training practice and would focus training research on ways to make experiential techniques as effective as they have proven in business and the military.

1. Variables. The independent variable is participation/nonparticipation in a program that employs experiential learning techniques. The dependent variables are learning immediately at the end of the program, retention of this knowledge, and job behavior and attitudes.

2. Treatment. An experiential program uses active experiential exercises to support and drive home most concepts presented through lecture and discussion. Exercises are generally group oriented and include simulation, role playing, group problem solving or decision-making, and gaming, to name a few. Use of these techniques is typically coordinated with substantive presentation on related principles and with debriefings on results.

3. Selection. The programs to be compared should be operated by the same organization, should be of comparable duration, and should use the same instructors to convey similar material to trainees with comparable background and experience. Ideally, trainees will be randomly assigned to conditions, so some receive experiential treatment and others receive traditional lecture methods. Because this may not be feasible, statistical equation of the two groups may be necessary.

4. Comparisons. Within- and between-group comparisons should be made (pre- and posttraining; experiential and analytic groups).

5. Measurement. In-program learning should be measured in three ways: pre- and posttraining administration of the course final exam, structured observation of trainee behavior in behavioral simulations at the end of the course, and content analysis of trainee Action Plans for grasp of central course concepts. Retention should be measured by a follow-up administration of the final exam three months after return to the job. Behavioral and attitudinal change should be measured by a three-month posttraining follow-up on the implementation of trainee Action Plans.

E. EFFECTS OF MAJOR NATIONAL PROGRAMS ON CAREER DEVELOPMENT

Police agencies look to major national police management training programs as providers of the highest quality training available. Claims of quality hinge on the belief that program graduates are uniformly successful in their later careers. Graduates are expected to experience an accelerated career development process, due partly to the quality of instruction received (the "value added") and partly to the reputation that these programs possess (the "ticket"). The actual effects of major national programs on graduate career development are undocumented. Do the careers of major national program graduates develop further and more rapidly than those of comparable nongraduates?

If graduates experience enhanced career development due to attendance at major national programs, they should receive more promotions, achieve higher terminal positions, and achieve terminal positions faster than nongraduates.

If this expectation is verified, then use of such programs to enhance the career development of promising managers can be partially justified. If not, the effectiveness of these programs should be examined in terms of other outcomes. The high cost of operating these programs would also require scrutiny. Even if research outcomes are positive, it will remain unclear how the "value added" and the "ticket" contribute to accelerated career development.

1. Variables. The independent variable is participation/nonparticipation in a major national police management training program. The dependent variables are promotions, highest terminal position, and time elapsed before reaching highest terminal position.

2. Treatment. There is no specific treatment that defines these programs. They tend to share some common characteristics: residential, extended duration, highly selective in accepting applicants, extremely costly to operate and/or participate in. The only intrinsic characteristic that these programs must display, however, is perception as among an elite group of programs with national prominence.

3. Selection. The program(s) selected should be among the elite group of six to eight programs with strong national reputations. Although it would be virtually impossible to accept and reject eligible applicants randomly, it may be possible to develop two generally comparable groups of program graduates and nongraduates by drawing on a list of eligible non-attendees for the latter group. It may be necessary to equate the two groups statistically on certain experience and background factors.

4. Comparisons. The post-program career development of graduate and non-attendee groups would be compared.

5. Measurement. Data on promotions, highest terminal position, and time elapsed before reaching highest terminal position should be obtained through a five- or ten-year follow-up survey of graduates and non-attendees. The survey should include questions on professional positions related to law enforcement but not on an operational level.

F. EFFECTS OF FORMALIZED SUPERIOR/TRAINEE PRETRAINING EXPECTATIONS UPON JOB BEHAVIOR

Trainees tend to make meager implementation effort once they return to their jobs. This is due partly to the lack of supervisor efforts to guide the trainee in formulation of pretraining expectations. As a result, trainees have little idea what their supervisors want them to learn in a course and implement later on the job. Consequently, trainees have little organizational motivation to learn, the organization appears uninterested and even hostile to what trainees manage to absorb, and trainees have little reason to try new behaviors. Were trainee and supervisor to formulate mutually accepted formal expectations for trainee learning and later implementation, would learning and successful implementation increase?

If documented and mutually accepted pretraining program and implementation expectations have their intended effects, trainees should learn more in a course, draw up Action Items that are likely to square with organizational interests, and implement new behaviors with greater success. Positive research outcomes would bolster the current outcry for job relevance not only in selection but also in how agencies prepare trainees for training and in how they receive them after training. This would indicate to LEAA, POSTs, and other funding/oversight agencies how to promote the usefulness of training through implementation of results. Negative research results, in contrast, would reaffirm current practice of providing little guidance to trainees in formulation of pretraining expectations.

1. Variables. The independent variable is pretraining formulation/nonformulation of mutually accepted trainee/supervisor learning and implementation expectations. Dependent variables are knowledge immediately at the end of the program and implementation of intentions formulated by trainee at the course's end.

2. Treatment. Pretraining learning and implementation expectations document course learning objectives and the areas and levels of later implementation as mutually agreed upon by trainee and supervisor.

3. Selection. Programs selected for study should have moderate orientation toward implementation. Trainees should be randomly assigned to conditions before training, with some required to formulate pretraining expectations and others not. (The requirement to formulate expectations should be directed through the chief to supervisor and trainee in parallel.)

4. Comparisons. Within- and between-group comparisons should be made (pre- and posttraining; with and without documented pretraining expectations).

5. Measurement. Learning should be measured on a pre-/posttraining administration of the course final exam and by content analysis of trainee Action Plans developed at the end of the course. Success in implementing new behaviors should be measured by a three-month posttraining Action Plan follow-up.

* * * *

The six research approaches outlined above do not exhaust those relevant to future policy and funding decisions. There are several other questions that should also be given a high research priority because police management training's future depends on the answers to these questions. The questions include:

- Do those who complete departmental preservice and refresher courses on their minimum duties and responsibilities display greater compliance with departmental policies and regulations?

- Do graduates meet with more success implementing new practices in smaller departments than in larger ones?

- If graduates of major national programs experience more rapid career development and achieve higher terminal positions than comparable nongraduates, does this result from the credential they have secured (the "ticket") or from their greater proficiency ("value added")?

- Does the use of management training as a tool in departmental decision-making result in smoother accomplishment of organizational change than the imposition of change from above, without rank-and-file input into the decision?

- What role does departmental training in MBO play in accomplishing department-wide changes in management philosophy?

- When departments have attained what we have called a "critical mass" of similarly well-trained managers, does large-scale organizational change necessarily follow? If not, what other factors intervene?

The next and final chapter of this report summarizes study findings and recommendations.

Nine: Conclusions And Recommendations

Throughout this report, we have drawn conclusions and made recommendations when we saw these were warranted by available evidence and by our observations and analyses. This chapter summarizes the conclusions and recommendations that are dealt with more fully in related chapters of the report.

A. CONCLUSIONS

Our conclusions fall into eight categories: the recent spread of police management training programs, program development practices, factors affecting program development, program models, program evaluability, single program evaluation approaches, evidence of program effectiveness, and national policy questions.

1. Recent Spread of Police Management Training Programs. Although reliable data about the recent multiplication of state and local police management training programs are unavailable, it is evident that their number has been on the rise. Why have their numbers risen so rapidly? Apparently, for several reasons. More and more people recognize the managerial shortcomings of the typical police manager. Beliefs about the desirable type of police manager have become more varied, requiring additional programs to reflect these beliefs. State and local authorities have demanded programs better geared to state and local needs and concerns. State and local programs have also come to be seen as less costly to operate than out-of-state, residential programs. The lower per-trainee cost permits more officers to be exposed to training. Officers attending training nearby can be called back in emergencies or to resolve coverage problems. If officers can commute to training, there is a lower likelihood of strain on family life. The recommendations at national commissions about expanded managerial training opportunities seem to have had some impact. POSTs have also broadened their influence and sought to strengthen their offerings. LEAA has made funds available for program development through SPAs and POSTS.

The spread of programs on the state and local levels may well be reversed in the near future, depending on whether and to what extent LEAA continues to support training, directly and through SPAs and POSTs. LEAA's probable withdrawal from police training support will have a critical impact on state and local training opportunities if the FBI simultaneously curtails its own training activities, as current budget proposals suggest is inevitable, and if the climate of fiscal austerity chokes off the appropriation of state funds for nonmandatory training programs.

2. Program Development Practices. Program managers and operators are generally familiar with the process for systematic training program development popularized by industry. This awareness has fostered expectations about how

programs should ideally be developed but has not solidified the field's estimation of the real value or feasibility of systematic program development.

How closely do police management training programs follow the industrial model of training program development? Our first reaction was to say that programs are developed "by the seat of the pants," hardly in a deliberate and systematic fashion that resembles the industrial model. For example, program developers and operators set goals largely without substantial input from user groups. They do little or no formal needs assessment. They typically skip over the identification of performance deficiencies and often pass off topical interest surveys as needs assessments. They do not set consistently clear objectives. They generally do not identify criteria that indicate, in measurable terms, the areas in which change is desired and the intended extent of change. They usually cannot pinpoint how they want in-program outcomes to affect job behavior. They rarely specify the learning principles that underlie instructional methods. They often fall back on an established curriculum long after its use has become counterproductive. They rarely provide a mechanism to help trainees and their superiors come to agreement about the individualized purposes for participation in training. They hire trainers based more on anticipated trainer-trainee rapport than on subject matter familiarity. They tend to exert little control over the composition of a class. They make minimal efforts to coordinate trainer activities. They distribute rewards for training almost indiscriminately. They obtain too little advance information about trainees for it to be useful in targeting content or measuring outcomes. They conduct few evaluations other than course critiques. They use evaluation results to tinker with program components but rarely to make needed major revisions.

When we stepped back and took a broader view, our reaction was somewhat different. Practices in development of police management training programs are not unreasoned. The process is deliberate, phased, and rather systematic. It corresponds inconsistently from point to point, however, with the prescribed steps of the industrial model. The practices we described only partially verify the claim by one program administrator that, "Programs are not rationally designed. Instead, they evolve--are gradually shaped by what is needed."

3. Factors Affecting Program Development. How systematic can program development reasonably hope to be? Our research confirms that there are factors external to programs and largely beyond the program developer's control that restrict how closely developmental practices can correspond to any chosen system. Expectations about how external factors limit development tend to be defined more clearly than expectations about what training is meant to accomplish and about how these ends will be met. Five external factors affect program development at each major developmental juncture. These are funding, legal requirements, organizational environment, community environment, and the ready availability of materials and resources from prior programs. Other factors affect program development at specific junctures. The homogeneity and stability of the target population, for example, can influence the feasibility and usefulness of a needs assessment. Departmental coverage requirements and the priority that user agencies place on training can dramatically affect availability of the type of trainee for whom a course was designed.

Our findings show how the program developer's options can be ringed by a multitude of forces beyond his control--forces that seemingly conspire to make the program developer abrogate his own assumptions. Faced with these constraints, most program managers and operators have rightly concluded that a high level of adherence to the industrial model is infeasible now and is unlikely to become more feasible in the near future. We stress, however, that most program managers and operators have the capacity to make isolated changes, for the interim, that will make their programs more manageable. But we also add that many of the obstacles to developing programs more systematically are not external to programs but are internal to their managers and operators. In response to a question on our national survey about needs assessment, roughly two-thirds of the respondents said candidly that they did not do a more formal needs assessment because program staff and user groups share such a close relationship and understanding on an informal basis that something more formal is unnecessary. Half of our respondents admitted that program staff believe that certain needs must be addressed, regardless of whether user groups happen to be conscious of them, so a more formal process would waste effort and could be counterproductive. Based on such remarks, it seems that a program developer's attitudes toward the value of systematic development can play an important role in activating and maintaining certain obstacles to more systematic practice.

4. Program Models. Do police management training programs follow a single model or several different ones? Our research shows that police management training programs take many forms, varying in the functions they seek to serve and in the means used to achieve chosen ends. They deliberately use widely differing types of resources, do different things with trainees, try to produce different types of changes in trainees during a course, expect and want trainees to try different types of things back on the job, and hope to impact on police agencies and the larger criminal justice system in different ways.

We found 14 variants of police management training in the field. Eight of these are basic models, each tied to a body of substantive information and reflecting the processes for transfer of knowledge along with related skills and attitudes. The eight show how training can be used to produce familiarity and compliance with departmental policy; to disseminate a prescribed body of knowledge derived from industry, the experiences of other police agencies or new laws; or to provide trainees with the concepts and experiences to manage more participatively. The other six are auxiliary models, which express no substantive content, must be tied to basic models to find substantive content, and focus on broader department- or system-level impacts and how to achieve them. The six show how training can be used to boost trainee and agency morale, to certify experienced and trained managers and weed out incompetent ones, to perpetuate the training experience beyond the classroom through an interactive network of course graduates, to recognize and anoint managers already tagged for promotion to the senior ranks, to facilitate two-way communications between senior departmental staff and line managers about a pending decision, or to build up the critical mass of managers similarly attuned to organizational change.

No single model was either fully articulated or unequivocally espoused by the programs we observed. In other words, none of the programs fully expressed any model, and all mixed several models together. The nature and causes of this

model mixing vary among and even within programs. Much of it is officially recognized, is set forth in public descriptions of programs, and fits together comfortably. But a lot of mixing also stems from lack of coordination within programs of curricula and instructional personnel and results in an unintended "smorgasbord" type of program that points trainees simultaneously in several directions and, ultimately, in no clear direction at all. A lot of model mixing also stems from the different roles and responsibilities that people have in a program. The problem is, this mix or "coexistence" among several models in a single program often produces ambiguity about the model or models in which the program is operating and, for that matter, those by which it ought to operate. As a result, people develop divergent notions about trainee selection, staff hiring, instruction coordination, compliance with state program requirements, program amenities, needs assessment procedures, curriculum design, and other matters. Thus, those with different functional responsibilities often do not act in concert. The variation among models and the phenomenon of model mixing may be viewed as legitimate, incidental, or just inevitable. Regardless, they have enormous implications for how programs should be managed and, as a corollary, how they should be evaluated.

5. Program Evaluability. How do programs stack up against the criteria for operation of an evaluable program, i.e., one in which resources can be effectively managed and an evaluation conducted with a reasonable chance of being useful? Even when the evaluation technology needed to determine program effectiveness happens to be available, programs typically fall far short of the evaluability criteria. For example, program expectations and activities are generally not at all well defined. Significant gaps and contradictions exist between program descriptions offered by policymakers, program managers, and program operators. Program expectations are often implausible, in light of general educational and training theory, the extent and types of resources brought to the program, the manner in which resources are used, and evidence of program relevance and effectiveness from prior program experience. Existing data collection systems to document program effectiveness typically rely primarily on trainee reactions and final examination scores, which do not provide necessary and sufficient information to show whether programs succeed in changing the trainee behavior on the job. Alternative data collection systems often impose unbearable cost or political burdens, involve unreliable tests or unstructured subjective observations by instructors, and anticipate uses of evaluation data that are beyond the control of program management. Based on the multiple roadblocks to evaluability that most programs have to contend with, we are forced to conclude that most police management training programs are far from optimally managed and that a major investment in evaluation is not what they really need.

6. Single Program Evaluation Approaches. The key question in program evaluation is not whether trainees liked a course or learned anything new from it but whether training made any difference in later job behavior. Few programs attempt to answer this question. There are many ways to answer it, ranging from a follow-up retest of the course exam for long-term retention to structured observation of work behavior. We examined three approaches that have wide applicability and require relatively low investment. The Action Plan Follow-Up Approach asks: Did trainees successfully carry out on their jobs the intentions

they took away from training? It involves systematic follow-up on implementation of personal action plans that trainees developed at the end of a course. The Proficiency Analysis Approach asks: Did training improve trainee proficiency in ways that the organization could use? It requires that trainees be tested before training for their managerial training needs and tested again a few months after training to see if training needs have been reduced. Simplified Cost-Benefit Analysis asks: Is training a reasonable investment to make, based on its return to the organization in increased job proficiency? It builds on Proficiency Analysis by adding a cost component that covers the costs to the organization of the manager's position, the costs of sending the manager away to training, and the amortization period over which the organization should benefit from training. All three evaluation approaches have shown their usefulness in industry and government, but none has been extensively used to evaluate police management training.

7. Evidence of Program Effectiveness. To what extent are police management training programs in general effective? The only evidence most programs can point to is reaction surveys and final course examinations. From these we can reach only two conclusions. First, trainees leave nearly all programs, regardless of variations in quality among them, satisfied that their time has not been wasted. Second, trainees leave programs with as much as or more knowledge of the rudimentary concepts dealt with during a course than they carried into it.

The evaluation literature offers scant evidence showing that the knowledge and skills trainees learn are relevant to their personal needs, will be usable in the organizations from which they come and to which they return, and has measurable payoff in greater individual proficiency and organizational effectiveness. The limited evaluation evidence seems contradictory at points and allows us to draw few conclusions about the overall effectiveness of police management training programs. It may well be inappropriate to expect an evaluation to give clear indications of effectiveness, however, because programs tend to be unclear in what they are trying to accomplish. As we concluded earlier, programs mix training models together, often legitimately but with the effect of having several partially implemented programs under one roof. This implies that it is altogether inappropriate to hope that the effectiveness of police management training in general could ever be demonstrated using a single set of indicators.

8. National Policy Questions. What are the important national policy questions in this area? There are a dozen or more such questions on which the future structure, practices, and orientation of police management training hinge.

- Do residential programs have greater effects on knowledge and proficiency than nonresidential programs?
- Are residential programs cost-effective in comparison with nonresidential programs?
- What effect does a network of program graduates have on the management practices, career development, and interdepartmental relations of network members?

- Do programs that use experiential teaching methods have more influence on later job behavior than programs that cover similar material but rely exclusively on traditional lecture-and-discussion methods?

- Do graduates of major national programs experience more rapid career development and achieve higher terminal positions than comparable nongraduates?

- Does the pretraining cooperative development by trainee and supervisor of personalized learning objectives result in higher subsequent implementation levels than traditional trainee preparation?

- Do those who complete departmental preservice and refresher courses on their minimum duties and responsibilties display greater compliance with departmental policies and regulations?

- Do graduates meet with more success in implementing new practices in smaller departments than in larger ones?

- If graduates of major national programs experience more rapid career development and achieve higher terminal positions than comparable nongraduates, does this result from the credential they have secured (the "ticket") or from their greater proficiency ("value added")?

- Does the use of management training as a tool in departmental decision-making result in smoother accomplishment of organizational change than the imposition of change from above, without rank-and-file input into the decision?

- What role does departmental training in MBO play in accomplishing department-wide changes in management philosophy?

- When departments have attained what we called a critical mass of similarly well-trained managers, does large-scale organizational change necessarily follow? If not, what other factors intervene?

B. RECOMMENDATIONS

Our recommendations for program management, operation, evaluation and utilization are directed toward three groups: user agencies, program operators, and program managers.

1. User Agencies. We recommend that user groups consider the following:

- In deciding whether to contract with outside organizations to deliver training, and in selecting trainees to be sent away for training, become familiar with the implicit and explicit objectives of available programs and determine whether these objectives are really congruent with the police agency's needs.

- Document the individual training needs that warrant sending a manager to training, communicate them to selected trainees and ensure that before training supervisors negotiate a set of personalized learning objectives with trainees and after training debrief them on training outcomes.

- Make greater efforts to measure the relative effectiveness of the programs to which the police agency sends its trainees at changing trainees in desirable ways. For example, review the intentions that trainees bring back from training and any demonstrated changes in trainee proficiency against agency needs. Maintain receptiveness to evidence that training has had a positive or negative effect on job behavior.

2. Program Operators. We recommend that program operators:

- Compare current developmental practice to the industrial model, identify steps at which current practice departs from the prescribed model, and analyze the benefits and associated costs of bringing practice into closer alignment with industrial standards. Direct particular attention to the feasibility and desirability of documenting the target audience's needs prior to program design and of later collecting evidence of training's effects on job behavior. Do not make isolated changes just to come closer to the industrial model, however. Take a complete inventory of the program as it presently exists first.

- Identify the external conditions that affect program development at each juncture of the process, assess their potency, identify the benefits of reducing their potency and the actions needed to do so, estimate the costs of these actions, and carry through on actions that promise valuable latitude in program development.

- Clarify the program's assumptions about the problems that give rise to the need for training and about the feasible solutions to these problems. Similarly, clarify what resources are needed to carry out program activities, how activities interrelate, what in-program outcomes should flow from these activities, how these outcomes relate to each other, and what long- and short-term impacts result from in-program outcomes. Where assumptions and expectations are unclear or conflicting, work to sharpen these aspects of the program model or bring them into closer alignment.

- Periodically examine how well the program meets the criteria for evaluability by completing the Evaluability Check list. Do not undertake a program performance evaluation without first assessing evaluability. If this self-assessment shows that a performance evaluation is not likely to be useful at present, then identify the adjustments in program expectations, activities, and information systems that might bring the program to a state of acceptable evaluability. Try to identify the full range of potential adjustments that might make the program more evaluable prior to undertaking particular adjustments.

- In program evaluations, focus where feasible on the central question: Does training make any difference in later job behavior? In designing a results-oriented evaluation, determine how the new data will relate to older data already being collected.

3. Program Managers: To program managers, including funding and oversight agencies, we recommend:

- Recognizing that program operators often perceive they have little to gain from following systematic program development procedures, examine the relevance of industrial standards to the programs under one's immediate jurisdiction. If this review shows systematic development would improve job behavior or program documentation at a reasonable cost, then develop incentives and other supports for programs that adhere to industrial standards. If program developers depart from the industrial model in unacceptable ways due to lack of knowledge or skill, then initiate efforts to familiarize them with appropriate standards and to transmit needed program development skills.

- Determine if funding or legal requirements inadvertently reinforce counterproductive program practices. Modify them if they do. As a result, curricula requirements should reflect a reasoned process and be regularly reviewed for continued relevance. Annual training requirements should be complemented and sharpened by documentation of the individual trainee's needs. Instructor standards should reflect ability to work within the endorsed program models. Training should focus on implementation rather than knowledge. Evaluation should focus on changes in job behavior.

- Clarify the training model or models that operating programs are supposed to follow.

- Cease the practice of rewarding trainees with certificates or monetary incentives for just sitting through a course and for perhaps also taking an examination. Try to get trainees to think of post-training implementation efforts as each course's final exercise. To foster this view of training, require trainees to report back to the program operator (or other appropriate person) on successes and problems faced in implementation efforts. Do not, however, tie reward decisions to the level of success in implementation.

- Encourage programs to view documentation of their relevance and effectiveness as an integral part of regular operations, but emphasize the periodic assessment of program evaluability. Conduct evaluability assessments through a combination of program self-assessments, state-level audits, and review by outside evaluators. Work with programs to improve their evaluability, with a long-range view toward termination of programs that do not make adequate progress toward an acceptable evaluability level.

- Promote the exchange of information among programs about program development practices, training strategies, and evaluation approaches.

- Direct research efforts toward issues that have major implications for later policy and funding decisions and that pose data collection requirements beyond the capabilities of individual programs.

DEFINITIONAL ISSUES AND EXPLICATION OF DEFINITION

Operationalization of the term "police management training program" required clarification of four key issues and resulted in a six-element definition that took in most programs self-defined as management training.

A. OPERATIONALIZATION ISSUES

Operationalizing the definition required resolution of four issues: restricting the word "police" and differentiating between training for managers and training related to management, between education and training, and between the view of program as client flow/delivery and a broader view of program that includes any activities performed to support the training situation.

1. Restriction of the Term "Police." "Police" potentially encompasses all forms of Federal, state, and local law enforcement; private security; certain support personnel; and so on. Even if private security and support personnel are excluded outright, all Federal, state, and local law enforcement agents remain. Some of them, however, are only tangentially involved in traditional police functions. Fire marshals and game wardens, for example, possess limited powers of arrest but perform few other enforcement functions. Federal police also serve specialized functions: enforcement of Federal laws, provision of police services to Federal buildings and properties, and provision of technical and training assistance to state and local law enforcement agencies. State law enforcement personnel focus largely on the restricted area of highway patrol. Municipal and county police agencies alone are engaged substantially in the traditional police functions of patrol, traffic control, and criminal investigation. County sheriffs serve dual enforcement and correctional functions. So, from a purely functional perspective, the term "police" could be restricted to municipal and county police agencies. The exclusion of state agencies, however, would be inappropriate for two reasons: the NEP's interest in both State and local activities and the common mixing of state and local personnel in training delivery. Thus, it seems appropriate to restrict "police" to include state and local police agencies performing police-related functions.

2. Training vs. Education. Any definition of training requires an at least implicit effort to differentiate training and education. None of the variables commonly used to differentiate training from education allow us to make clear distinctions. Some variables obscure more than they elucidate, for example, psychomotor vs. cognitive activity, practical vs. theoretical orientation, restricts alternatives vs. opens up alternatives, department- vs. self-financed, department- vs. self-initiated, department vs. self as principal beneficiary, short vs. long duration, accelerated vs. distributed, and so on. The concept "management training" can in itself be seen as a contradiction in terms, if training is considered to be narrowly focused on outlining specific steps in a task, whereas management broadly involves the coordination of many

CONTINUED

inexplicable variables. Even the differentiation between training and education as practical vs. theoretical fails to hold up and merely highlights two polar views of what training ought to accomplish. The former views training narrowly as vocational preparation to perform a given checklist of tasks and downplays the trainee's understanding of the interrelationships among tasks, how roles and functions mesh with other criminal justice and social service agencies, and what the police manager's role ought to be. The latter views training in the broadest perspective as providing understanding of current and alternative roles and responsibilities and intermeshing of roles with other agencies. Most programs claiming to offer police management training candidly admit, "We are doing far more education than training." The exception is programs geared to gain compliance with a particular agency's policies and regulations. Moreover, the differences in intent between training and education requirements are often also unclear. As a result, a police manager in some areas of the country may obtain either training or education credit for the same instructional sequence. The emphasis of many training administrators on assuring that an individual does not receive double credit highlights the lack of clarity, on a day-to-day basis, between the results expected to flow from training vs. education. Consequently, any distinction we draw between the two will necessarily be an arbitrary one based on technical criteria.

3. Training for Managers vs. Training Related to Management. The label "management training" is applied undeservedly to many programs that are not related to management and supervision but are offered to persons who are managers. Many supervisory courses, for example, contain much "refresher" material, unrelated to the supervisory responsibilities of participants. In practical terms, management training often refers to "any training offered to managers." Some program managers suggest that the term refers to "whatever is hot, the latest fad," or "whatever someone wants it to mean at the moment." Management training will, thus, be defined more strictly in terms of a process to improve the performance of managers in using resources, influencing action, and facilitating change to accomplish organizational goals.

4. Program as Delivery vs. Developmental Process. The term "program" may be interpreted narrowly, as the process of involving trainees in a series of activities, or broadly in terms of all activities related to preparation for, support to, and assessment of the training situation. Both interpretations are valid, though the former is more consistent with prior NEP studies, and both are used in the study, although the emphasis will be on the narrower definition.

B. ASSUMPTIONS BEHIND THE DEFINITION

To permit operationalization of the topic area definition, we made several assumptions based on the distinctions outlined above, including these:

- Although a distinction should be drawn between training and education, training programs cannot be considered education merely because they happen to occur in an academic environment. One practical distinction between training and education reflects the formal relationships of trainees to an academic institution: if the trainee is required to accept academic credit for a course or to gain formal admission to a school before participation is permitted, then a program otherwise regarded as training should be considered education.

- Although the study focuses on efforts by state and local agencies to operate programs, national and regional programs that serve state and local agencies should not be excluded, including those operated by Federal agencies, such as the FBI.

- Only programs serving state and local law enforcement agencies should be included.

- Many training courses for police managers include nonpolice personnel. Some latitude should be given to include these programs.

- All levels of police management training should be included, from first line supervisor to chief executive.

- Included programs should aim at improving one or more aspects of management performance, not refreshing the "basics" of police work.

- The ultimate beneficiary of intended performance improvement should be the police agency, not the individual trainee.

C. THE OPERATIONAL DEFINITION

For the purposes of this study, the term "police management training program" was defined as: an instructional sequence of relatively short duration, offered outside a degreed academic program for the purposes of upgrading one or more aspects of supervisory or management performance for the ultimate benefit of a police agency to currently active or soon-to-be-commissioned police supervisors or managers who operate at the state or local levels. To be regarded as a police management training program, a given program had to meet all six criteria, each of which specifically includes and excludes certain programs.

1. An Instructional Sequence. This includes courses of instruction that remove the individual from the regular work routine. Excluded are internships, job exchanges, roll-call training, correspondence courses, and other educational activities that might be directed toward similar objectives but that are incorporated into the work routine.

2. Offered outside a Degreed Academic Program. This excludes programs that require the trainee to make formal application to the host academic institution and those that require the trainee to accept academic credit instead of offering credit as an option. All other programs are included.

3. For the Purposes of Upgrading One or More Aspects of Supervisory or Management Performance. Included are programs geared to effect change in administrative, supervisory, or management performance by any means, including capacity building, technology transfer, attitude change, introduction to participative management principles, and so on. Also included are instructional sequences that focus primarily on improvement of performance in management areas but that include a small amount of specialized and refresher material. Excluded are programs that are primarily operations rather than management oriented, focusing on specialized police practices or refreshing skills learned in basic training.

4. For the Ultimate Benefit of the Policy Agency. This includes all programs taken in compliance with departmental or state training requirements, programs otherwise taken at the behest of a policy agency, and programs recognized by the agency as job related. This excludes programs taken for personal development that are of no direct benefit to the agency.

5. To Currently Active or Soon-to-Be-Commissioned Police Supervisors or Managers. This includes programs that are offered to any combination of current or soon-to-be-commissioned police supervisors and managers, from officers under consideration for promotion to a first line supervisory rank, up to the department's chief executive. Also included are programs offered to representatives of criminal justice agencies other than police, where this is either a regular economy measure or a deliberate effort to increase inter-agency communication and effectiveness. This excludes courses offered to a general audience that occasionally includes some police managers.

6. Who Operate on the State or Local Levels. Included are programs offered primarily to state and local police agencies performing substantial law enforcement functions. Excluded are programs attended primarily by Federal law enforcement personnel.

The definition above reflects an interpretation of "program" centered around training delivery. For the definition to apply to the developmental interpretation of program, this clause would need to be added: as well as to all activities necessary to plan, operate, support, and assess the training situation.

METHODOLOGY

The NEP approach sets forth a phased, closely reasoned descriptive and analytic process to provide the basis for proceeding with an evaluation likely to prove useful. Thus, it differs from more traditional approaches because program evaluations are typically planned and implemented in isolation from decision-maker interests and operating program realities and often produce at great expense information that is of little or no use to program managers and operators.

In this appendix, we discuss the study methodology in three parts. First, we lay bare the assumptions behind the NEP approach. Second, we summarize the work tasks followed by all NEP studies. Third, we specify our data sources and data collection procedures.

A. ASSUMPTIONS BEHIND THE NEP APPROACH TO EVALUATION

The NEP evaluation approach builds on clear assumptions about the rationale for a phased data-acquisition strategy, the conditions needed for a program to be evaluable, the evaluator's role in working with program managers to create these conditions, and the value of this process for making programs work and demonstrating that they work.

1. Rationale for a Phased Data-Acquisition Strategy. Because nearly any type of program will, at any point in time, be in a state of only partial implementation and because a program's operators will generally experience ambiguity in expectations and in direction from program managers and policy-makers, it is inadequate to base an evaluation strategy upon preconceived concepts of how and toward what ends a program will operate. For a program to be usefully evaluated, there must be a commonality, or match, between what those in charge of a program believe it to be and what actually exists. Because the evaluator cannot know in advance the degree of program implementation and the level of correspondence between the rhetorical program expectations shaped in the political world and actual day-to-day operations, it is fitting to adopt a phased evaluation approach.

The evaluator, thus, has the initial task of seeing if the expectations held and the questions being asked are in reasonable alignment with the actual program as it operates on a mundane, everyday basis. The assumption here is that, if we rush headlong into evaluation before certain preliminary matters are addressed, we run the risk of evaluating a rhetorical program that does not exist in reality and of, thus, generating information of little or no use to policy-makers or program managers. On the other hand, if we take the time first to answer preliminary questions about program evaluability, we often find it is not necessary, or not immediately feasible, to conduct a full-scale evaluation.

Nonetheless, through this process, we may find enough information to answer our initial questions and may want to redesign aspects of the program, modify program goals and objectives, or set up a monitoring system before entertaining a full evaluation. Evaluability assessment can also help justify freeing up resources to conduct a full evaluation because it anticipates arguments for not evaluating by establishing that an evaluation will be useful. In such circumstances, it provides a clear notion of what should be evaluated. It does this by identifying performance measures that are mutually acceptable to policymakers and program operators, validly represent program activities, are plausible in light of program resource allocation and general theory, do not impose unrealistic cost or political burdens, can reliably represent the phenomenon under study, and are likely to be used to manage the program better.

2. Conditions Needed for a Program to Be Evaluable. A program must meet seven evaluability criteria for it to be likely that its evaluation will prove useful:

- Program management defines with reasonable completeness what is expected to happen in and result from the program.
- Management's intended program corresponds with policymakers' expectations and is, hence, acceptable to them.
- The program actually in place validly represents program management's expectations.
- The program's expectations are plausible, in light of accepted general theory, allocation and utilization of program resources, and the evidence from the program's experience.
- The program's intended data collection procedures are feasible, in that they pose no unacceptable cost or political burdens.
- The program's intended data collection procedures are capable of generating reliable information.
- The program's intended uses of evaluation information are plausible, in that they are under the program manager's control.

3. Evaluator's Role in Creating Conditions for Evaluability. The evaluator's role in the evaluability assessment process depends on the time and other resources available and on the evaluator's relationship to management. At a minimum, the evaluator has the responsibility to describe the program as it actually operates. This description should cover factors that limit its evaluability in terms of the seven criteria. The evaluator may also take on broader roles, as noted in the introduction. These other roles include assistance in the development of strategies for removal of roadblocks to evaluability, monitoring of the implementation of these strategies, and design of traditional evaluations.

4. Value of Evaluability Assessment for Program Management. The evaluator's close working relationship with program management in the identification and removal of roadblocks to evaluability, and, thus, in the design of a more evaluable program, shows the link between evaluability and manageability. A program that does not meet the criteria for evaluability cannot work with optimal effectiveness and/or demonstrate how well it works. Such a program is neither fully manageable nor fully able to maintain accountability. Evaluability assessment is valuable because, if it becomes evident that evaluation results will not be properly used to control a program, the evaluation effort is broken off before resources that could be better spent in reexamining and altering a program are misspent. By removing the roadblocks to evaluability, it helps programs to work and to demonstrate their successes.

B. THE NEP WORK TASKS

The study followed the generic task structure to which all NEP Phase I studies have more or less adhered. Although LEAA has revised and refined the NEP work tasks several times since the first NEP in 1974, the basic NEP structure has remained the same. To guide the study, we drew heavily on the latest formal revision of the NEP work plan, dated February 1977, and on other published documents on evaluability assessment. The latest NEP work plan calls on the evaluator to perform iteratively nine interrelated tasks, which we paraphrase as follows:

1. Determine the key policy issues and prevailing expectations for a program by review of enabling legislation, hearings reports, program brochures, grant applications, feasibility studies, master plans, progress reports, expert opinion, speeches by program officials, and other materials.
2. Compile past findings of fact for the program, such as evaluation studies, monitoring data, and program reviews; discuss potentially quantifiable data with program operators, evaluators, auditors, and managers.
3. Develop criteria for inclusion of programs in the topic area, test identified programs against these criteria, and survey included programs.
4. Anchor expectations about programs firmly in reality by visits to operating programs and subsequent description of them in narrative and graphic form.
5. Based on activities observed in the field and discussions with those having responsibility for program implementation, inductively derive general program models that show the logic by which operating programs presume inputs link to processes and processes to outcomes and impacts.
6. Synthesize the program models into a single measurement model that shows for key variables one or more measures that could be taken for purposes of program monitoring or evaluation.

7. Array at each appropriate measurement point the available data related to the program models, assess the conclusions to be drawn about program effectiveness, and identify gaps in the current state of knowledge about program effectiveness.

8. Develop approaches for obtaining more information by adapting the measurement model into evaluation designs suitable for use by single programs and by designing alternate research strategies for filling gaps in the state of knowledge about programs nationally.

9. Pretest and fine tune the evaluation designs for use by single programs and/or the research designs for filling gaps in knowledge, then feed the pretest results back into earlier tasks.

The NEP work plan specifies nine work tasks but explains that they cannot be performed sequentially. It stresses that tasks are interrelated and that the required process is iterative, calling for "successive correction" and repeated adjustments in topic area definitions, the universe of topic area programs, program models, performance measures, and suggested field data collection procedures. With increased knowledge and repeated adjustments, the topic area takes on clearer definition, the pool of candidate programs is iteratively reduced, program models become more refined yet better anchored in reality, and evaluation strategies become more cost-feasible yet more attuned to real concerns. In short, each task develops information that is used to sharpen the questions asked and to refine the products of other tasks.

C. DATA SOURCES AND DATA COLLECTION PROCEDURES

Precisely because the NEP methodology is phased and interative, it would be difficult and possibly unproductive to detail the interrelationships and movements among tasks that evolved in this _particular_ NEP study. The methodology requires data collection procedures that are much more complex and less straightforward than a traditional evaluation might call for. Prior efforts to describe this process graphically through flow charts have looked like labyrinths. Lines meant to indicate movement among tasks point back and forth between the same boxes, loop back repeatedly to earlier tasks, and defy identification of single starting and ending points. And what happens in one NEP study can be quite different from what happens in another, properly so because the process encourages flexibility in obtaining answers to questions that become clear enough to explore only part way through a study. We stress that, because tasks cannot be performed sequentially, it is not worthwhile to detail the sequence of tasks, but, in Exhibit 2-A, we present a rough list of basic data collection and analytical tasks in their order of performance. Certain underlying tasks occurred almost continuously, however, and cannot be shown.

We relied on five data sources for the raw material used to conduct our analyses. The sources were: consultation with experts, literature review and analysis, site visits, large-scale mail survey, and pretests of single program evaluation designs. (Each relates to particular NEP tasks, indicated by the numbers in parentheses in _descending_ order of importance.)

EXHIBIT 2-A

NEP/Police Management Training

LIST OF TASKS PERFORMED IN THIS STUDY

TASK	PURPOSE
Conduct preliminary review of written program descriptions, related policy statements, and other general literature	Identify key policy issues in program development and management, identify prevailing expectations, identify experts
Conduct preliminary review of literature related to evaluation and measurement	Determine typical measurement concerns and options, identify other sources of evaluation evidence, identify experts
Draft preliminary description of topic area	Provide framework for further data collection
Conduct preliminary surveys of POSTs, SPAs and selected programs	Locate programs, ensure that programs exist, clarify program expectations, determine practitioner concerns and interests
Refine/operationalize definition into discrete criteria	Identify candidate programs for further study
Determine whether programs fit definition	Finalize topic area universe
Draft site-visit guide	Ensure that interviews at each site address certain common issues yet are tailored to interviewee's vantage point and study's information needs
Determine criteria for selection of sites to be visited	Ensure that all visited programs share key properties but are reasonably representative
Select sites to be visited	Ensure representativeness before starting visits
Determine official expectations for each program before making visits	Determine how officials suppose that actual programs work, toward what ends, and so on
Conduct site visits	See how the program actually works so expectations can be better anchored in reality
Develop data files for each site	Gather raw data for completing the inductive analytic tasks
Draft analytic and evaluative reports that cover the study's basic purposes, based largely on site visits	Provide guidance for program management and research
Disseminate draft report for review by a wide range of potential users	Determine how to refine the draft report best to meet user needs and interests
Design and administer large-scale mail survey on selected subjects found in the draft reports	Determine if findings generalize to a broader universe of programs
Pretest evaluation design recommended for single program use	Fine tune evaluation designs so they may be adopted more readily; provide evidence these approaches are usable
Revise analytic reports	Better meet user needs and interests; incorporate results of survey and pretest
Disseminate results	Ensure that findings have their appropriate impact on program management and evaluation

1. Consultation with Experts (Tasks 1, 2, 3, 8, 5-7, 9). Expert consultants gave us a starting point and original sense of direction but also served as a sounding board throughout the study.

a. Preliminary telephone survey of POSTs, SPAs, and selected programs (Tasks 1, 3, 2). Early in the study, to locate programs, to determine what they expect to accomplish, and to sensitize the study team to practitioner concerns, we conducted a semi-structured telephone survey of POSTs in the 47 states belonging to the National Association of State Directors of Law Enforcement Training (NASDLET). These surveys covered topics like: the POST's role in management training; locations of the POST's management training programs and of other noteworthy programs in the state; the purpose for the POST's management training programs; availability of program performance data; techniques used in needs assessment, pretesting, and evaluation; environmental factors affecting program development and utilization of training; and the most important questions for this study to address. Many POST interviews were followed up by similar interviews with SPAs and with a broad sample of the identified programs. This allowed us to clarify program expectations further and to ensure that identified programs actually existed. In a prior report, the Evaluation Design Plan, we described the influence of practitioner interests on the study's overall design and conduct.

b. Ad hoc expert consultation (Tasks 8, 6, all others). Intermittently throughout the study, experts--in training program development, management training evaluation, and program management in particular--helped guide the study. There were many purposes for seeking and accepting expert opinions. These included sharpening our findings, finding means for prompt dissemination, and gaining technical advice in specialized areas. This assistance had greatest impact on measurement-related tasks. Staff of the U.S. Office of Personnel Management, Productivity Research and Evaluation Division, provided substantial input to measurement-related tasks. Several POST directors and the operators of independent programs also gave periodic assistance.

c. Report review (Tasks 5-8). After completion of site visits (Task 4) and analysis of site visit files in light of other information (Tasks 5-8) we disseminated a draft report for review by a wide range of potential users. The selected reviewers worked in LEAA, other Federal agencies (FBI, OPM), POST councils, professional organizations, and visited programs. They included program managers, program operators, evaluators, and researchers. Review comments were synthesized chapter by chapter to indicate how each chapter might be revised better to meet needs of two audiences: one composed predominately of researchers/evaluators and the other consisting mainly of practioners. This synthesis resulted in Technical and Summary Report outlines that were closely followed in development of study products.

2. Literature Review And Analysis (Tasks 1, 2, 7, 8, 6, others). Like expert consultation, the literature review spanned the entire study but served different purposes at different times. At first, it contributed to the original framework of prevailing expectations and concerns. Later, it served as a basis for comparison with on-site observations, as the source of evaluation measures and approaches, and as the raw material for assessing the state of knowledge about program effectiveness.

a. Initial literature review (Tasks 1, 2). The initial literature review covered program brochures and course descriptions, national and state reports on the quality of and need for management training, journal articles that outlined program strategies, a few annual reports from major programs, and several completed evaluation studies. The purposes for this initial review were similar in many respects to the purposes for early consultation with experts: to establish a sense of direction, to determine what expectations for police management training prevailed, and to identify burning issues. Another purpose was to get a sense for the range of current evaluation options and to identify the key figures in program evaluation.

b. Ad hoc literature review (All tasks). The literature review continued throughout the study. The reviewed literature included materials collected on site (program brochures, course syllabi, handout materials, attendance records, grant proposals, research surveys, evaluations, and unpublished articles), those mailed in by non-visited programs (mainly master plans, course syllabi, and program schedules), and those contributed by program evaluators (evaluations, journal articles, evaluation approaches). These materials were used in all the study's analytic tasks.

c. Analysis of completed evaluations (Tasks 7 and 8). After we developed program models and identified potential performance measures, all collected program evaluations were systematically reviewed. This constituted the formal literature review but was only one aspect of the continuing examination of program effectiveness.

3. Site Visits (Tasks 4 and 5). The on-site opportunity to observe programs, interview all interested parties, and review relevant documentation was central to the study's inductive approach. In particular, it provided the basis for empirical development of program models and accurate identification of roadblocks to evaluability.

a. Development of site-visit guide (Task 4). To ensure that interviews at each site addressed certain common issues but were tailored to the individual interviewee's vantage point and the study's informational needs, we drafted a 145-question site-visit guide. The guide was not designed to produce quantitative data but to develop comparable descriptions of individual programs. The guide consisted of a pool of issues from which discussion topics might be drawn during a given interview or observation period. Originally derived from general literature, expert opinion, and the telephone surveys described above, the guide was refined and expanded as site work progressed and additional issues were identified. In light of the exploratory nature of the NEP, the guide necessarily consisted mostly of open-ended questions. In some questions, response categories were indicated as potential probes to be used when warranted. The guide is contained in another report, the Evaluation Design Plan.

b. Site selection (Task 4). Because the topic area universe includes diverse programs, we formulated site selection criteria to ensure that all visited programs shared key properties but were also reasonably representative along certain dimensions. To be eligible for a site visit, a program had to meet these minimum criteria: corresponds with the topic area definition, was first conducted a year before the intended visit, program director expresses

interest in cooperating with the NEP study, program staff will be available on site for interviews, offers the opportunity to observe the training process directly, and is under 200 hours in duration. In addition, in selecting sites, we tried to balance them on certain variables: the level of training (supervisory, mid-management, executive development), auspices or sponsoring agency (SPA, POST council, academic institution, municipal agency, regional academy, state academy, FBI, national associations), geographic location, and predominant expectations. The study team conducted site visits to 16 programs. A list of visited programs is contained in Appendix 3.

c. Site-visit procedures (Task 4). One or two study team members visited 16 programs for an average of three days per program. Where feasible, we followed this sequence in collecting data from each program: advance telephone interview with program director and relevant oversight agencies to establish official program expectations; advance review of enabling legislation, program brochures, evaluation reports, and other documents; initial on-site interviews with program director to explain the study, clarify the program's expectations, and develop a site-visit schedule; direct observation of classes, interspersed with instructor interviews and interviews with curriculum developers and auditors; periodic briefing and querying of the program director and other core program staff; trainee interviews; review of program files and performance data; closeout interviews with program director and core program staff to reconcile discrepant bits of information; follow-up calls with key individuals unavailable on site and with program directors and appropriate oversight agencies to fill gaps in information collected on site.

As often as not, once on site, our data collection procedures bore little resemblance to this intended sequence. In some instances, a visit began with direct observation, followed by program director interviews at the end of the first or start of the second day. In two programs, situational factors prevented our observation of training activities altogether. (In one case, trainees were dismissed to the library for two days; in another case, the initial decision to cancel the program due to inclement weather was reversed when several trainees arrived.) Typically, half to two-thirds of our time on site was spent in direct observation of training. Interviews with program staff occurred when staff were available, often between classes. Trainees were interviewed once they completed the initial warm-up period of breaking down intra-group communications barriers and were comfortable with our presence. In five cases, we made site visits to the relevant oversight agency to complement data collected from the program. The site-visit guide was used to determine gaps in the information collected at a given program and to suggest ways for structuring particular interviews.

d. Site-visit files (Task 4). We developed a site-visit file for each observed program and one unobserved program. A file included information in seven categories: data management information, program profile, origins and history, program development practices, factors affecting program development, client flow models, and evidence of program effectiveness. Appendix 4 outlines the key file elements. The most important parts of a file are the last two. The section on client flow models describes in narrative and graphic form the intended and observed programs, analyzes disparities within and between them,

and identifies roadblocks to evaluability. The section on evidence of effectiveness assesses the quality of available program performance data and summarizes key unanswered theoretical and policy-related questions. We submitted these descriptive and analytic field notes in draft to the directors of visited programs to identify errors of fact and imbalances in perspective. Nearly all program directors found the reports substantially accurate, with some modifications, and considered them from somewhat to highly useful. Corrected reports were distributed to each program. The information contained in the corrected site-visit files constituted the primary data source for Chapters Two through Six.

4. Large-Scale Mail Survey (Tasks 4, 3). To provide a greater sense of breadth to certain findings, it seemed important to develop the basis for generalization to a wider universe of programs. To accomplish this, we designed and implemented a mail survey that extends the generalizability of the rich data collected on site.

a. Survey design (Tasks 4, 3). In conjunction with LEAA staff, we concluded that the survey should focus on areas in which practioners had expressed most interest: program development and factors affecting program development. From the draft findings, we selected items that generally called for objective reporting of facts and minimal projection of personal attitudes. About one-fifth of the items solicited demographic information about programs, one-half dealt with program development, and one-third dealt with factors affecting program development. Questions asked for three types of responses: check the one appropriate response, check all appropriate responses, rate all responses on a scale from 0 through 4. We pretested this draft instrument on seven program managers and operators and revised questions to reflect their feedback. Appendix 5 contains the Large-Scale Mail Survey.

b. Survey universe (Tasks 4, 3). Programs for inclusion in the mail survey were identified through several means. We drew on the preliminary POST/SPA surveys, on other information obtained inadvertently during the study, and on two special locator surveys conducted just prior to the mail survey. The locator survey, directed to POSTs and all police departments servicing cities with a population over 100,000, sought to identify programs' key personnel. Nearly all programs identified prior to the mailing of the large-scale survey were included in the survey sample. The few exceptions were those conducted by agencies from which two programs had already been included in the sample.

c. Survey implementation (Tasks 4, 5). We sent out each survey in a government-franked envelope, accompanied by a cover letter signed by an LEAA official and a franked return envelope. Two weeks later (a week before the due date), we sent reminders to the 80 percent that had not yet responded. Agencies that were sent more than one survey but had responded to only one were sent a special reminder. Two weeks later, a week after the due date, we sent similar second reminders along with second copies of the survey and set a new deadline three weeks away. One week later, we sent out another special reminder to programs that responded to one but not all surveys, to ask whether the responding agency intended its response to cover all of its programs. Meantime, the training bureau chiefs in two States, California and Florida, helped boost response rates by sending out an endorsement of the survey to all in-state operators of

relevant programs. Completed surveys were accepted for two weeks beyond the revised due date. Approximately 92 percent of the 250 sampled programs responded, including those reporting that they either had not offered a program in the prior year or did not fit the topic area definition. If we regard these programs as inactive and, hence, outside the universe and then subtract them from the 250, we achieved an adjusted completion rate of approximately 90 percent (183 of 206). This response rate was 10 percent higher than projected. Exhibit 2-B profiles the characteristics of programs that completed the survey.

d. Survey analysis (Tasks 4, 5). We calculated frequency distributions for all 304 variables contained in the survey. Because some respondents did not provide valid responses on some questions, we calculated adjusted frequencies that reflect only valid responses. (Most invalid responses were on rate questions, which were either left blank or checked.) Within questions that asked the respondent to check appropriate responses, we ranked the adjusted frequencies for relevant variables.

For variables that asked respondents to give a rating from 0 through 4, we also calculated a mean rating and the adjusted cumulative frequency for non-0 ratings (the sum of adjusted frequencies for 1 through 4 ratings). Within rating questions, for relevant variables, we ranked both the mean ratings and the adjusted cumulative frequencies for non-0 ratings. In calculating the means, we recognized that this was not technically appropriate because the mean presumes the intervals between data points are equal, which could not be said of our data. At the same time, we recognized that use of the mean in this circumstance is common practice and was the only way to reduce rating data to a common denominator meaningful to most readers. The analyses described above were performed on the entire sample and also for five subsamples to distinguish by trainee rank and jurisdiction. The five subsamples were: supervisors; mid-managers; executives; single municipality, county, or metropolitan areas; interstate region, statewide, multi-state, or national.

Resulting data were incorporated within chapters on program development and factors affecting program development to the extent that they enhanced or helped refine the original findings. We did not try to provide full tabular presentation of data because the survey's purpose was to let us sharpen and extend the generalizability of site-visit findings. We were careful that survey results did not displace site-visit findings because, in many respects, we place greater reliance on the observational data. The data collected on site were repeatedly verified through multiple interviews. In contrast, survey data were provided by single respondents, without verification, often in response to questions that imposed "demand characteristics," i.e., they asked for actual practices but strongly suggested that certain responses were relatively more desirable. On a few variables, where demand characteristics were particularly strong (use of task analysis; instructor selection practices), survey and site visit findings clashed but could be explained.

5. Pretests of Single Program Evaluation Designs (Task 9). To establish that the evaluation designs proposed in Chapter Six for use by single programs could produce useful information, and to fine tune these designs so they could readily be adopted, we pretested two of the three approaches. We deferred

EXHIBIT 2-B

NEP/Police Management Training

CHARACTERISTICS OF PROGRAMS SURVEYED BY MAIL

	Number	Percent
LEVEL OF PRIMARY INTENDED TARGET AUDIENCE		
First-line supervisors	98	53.0
Middle managers	46	24.9
Executives	27	14.6
Cross section of two above	8	4.3
Cross section of all three	17	9.2
JURISDICTION OF PRIMARY AUDIENCE		
Single municipality	32	17.3
Single county	18	9.7
Single metropolitan area	22	11.9
Intrastate region	19	10.3
Statewide	86	46.5
Multistate	5	2.7
National	3	1.6
TIMES OFFERED IN THE PAST YEAR		
Once	76	41.1
Twice	46	24.9
Three times	23	12.4
Four to six times	27	14.6
Seven to ten times	7	3.8
More than ten times	6	3.2
NUMBER OF TRAINEES IN AVERAGE SESSION		
1-5	3	1.6
6-10	3	1.6
11-15	18	9.7
16-20	40	21.6
21-30	59	31.9
31-40	25	13.5
More than 40	37	20.0
AUSPICES*		
State Planning Agency (SPA)	12	6.5
Police Standards and Training Agency (POST)	95	51.4
Academic institution	50	27.0
Municipal or county academy	36	19.5
Regional academy	26	14.1
State academy	38	20.5
FBI	5	2.7
State associations	10	5.4
National associations	4	2.2

* Multiple responses bring totals over 100%

experimentation with the third design because it builds on the second, which needed to be pretested first. Only programs with which we had reasonable familiarity were eligible for selection as pretest sites.

a. Action Plan Follow-Up Approach. The Action Plan Follow-Up Approach was pretested at two programs: an off-campus, two-week course conducted by Pennsylvania State University's POLEX program and a two-week course conducted by the Southeast Florida Institute of Criminal Justice at Miami-Dade Junior College. Prior to each pretest, we visited the cooperating program to negotiate the questions to be asked in the evaluation, the procedures for incorporation of the action plan exercise into the course, and the procedures for later collection of follow-up data. We formalized the negotiation process in the second pretest through use of an "action plan menu," which detailed preliminary questions and then set out the procedural options in two columns. Column A indicated required procedures, and Column B contained optional procedures. We negotiated listed options with the program director, circled those he chose, and added additional options that emerged during negotiations. From this process resulted a document that fully described all steps agreed upon for implementing the action plan exercise and in preparing trainees for the follow-up. The study team was present at the start and conclusion of both pretested programs, i.e., both for introduction of the action plan concept and for completion of the action plan exercise. Appendix 7 contains the forms actually used in conduct of both the action plan exercise and follow-up interviews.

Follow-up interviews in the two pretests followed the prescribed format and were generally similar. They differed in two key respects: scheduling procedures and time of day for follow-up interviews. Both differences resulted from the high percentage of trainees in the second pretest who worked nights. In the first pretest, we set up interviews with trainees by telephone one week before the week scheduled for interviews. We interviewed 17 of 19 trainees, most of them during conventional business hours. In the second pretest, we relied primarily on a mailed scheduling form that trainees were asked to back one week before the interview week. To compensate for lack of telephone contact with the second group during the week prior to the interviews, we sent xerox copies of trainee action plans along with the scheduling form and also verified scheduled interviews by letter or phone after the trainees returned scheduling forms. We interviewed 13 of 17 trainees in the second pretest, many of them outside conventional business hours.

Both pretests produced lengthy program evaluations. We presented results to the host programs in two ways. First, we mailed them to the cooperating programs for initial review. After the initial review, we made oral presentations of evaluation findings on site. The purposes of the oral presentation were to disseminate results, to determine how results might be used, and to identify ways to make the evaluation approach more useful. Program feedback was then incorporated into the finalized description of the approach in Chapter Six.

b. Proficiency analysis. The proficiency analysis approach was pretested at one program, a four-week course offered at Northwestern University's Traffic Institute. Here too, prior to the pretest, we visited the cooperating program to negotiate procedures and mutual assurances. The study team was present at the start of the course to explain the evaluation process and to obtain

pretraining self-assessments of training need on the Managerial Training Needs Profile (MTNP). During the course, through the program director, we sent a copy of the MTNP to each trainee's immediate supervisor to obtain the supervisor's assessment of the trainee's needs. Of 31 trainees, 22 filled out the MTNP at least partially; 11 of 27 supervisors completed the MTNP. Two months after the course's completion, again through the program director, we sent a second copy of the MTNP to each trainee and supervisor. The personalized cover letter accompanying each MTNP recognized the potential respondent as trainee or supervisor and as an individual who previously did or did not complete the survey. Two weeks after we sent out the post-training MTNPs (one week before the due date), we sent reminders to all nonrespondents. We also sent second copies of the MTNP to nonrespondents who had responded to the pre-training survey. This time, 14 trainees and 15 supervisors returned surveys. We analyzed pre- and post-training data to determine whether the course was improving proficiency (i.e., decreasing training needs) in the areas that trainees and their supervisors considered priorities. Because of low response rates, we had to abandon plans for multivarate analysis, with two points of view at two points in time. Instead, we conducted three other forms of analysis. To determine if trainees and their supervisors perceive similar training needs, we compared matched trainee-supervisor data sets. To assess training's effects on managerial proficiency, we first compared pre- and post-training group scores for trainees and supervisors. We then refined our analysis by looking only at cases for whom we had both pre- and post-training data.

The proficiency analysis pretest generated a brief report. As in the other pretests, we mailed the report to the program for initial review and later made an oral presentation on site. The program also sent a synopsis of the evaluation findings, which we drafted, to trainees and supervisors who requested evaluation results. Program feedback was then incorporated into the finalized description of the approach in Chapter Six.

Despite a few isolated problems with access to programs or data and in interpretation of findings, the study's methodology successfully incorporated the latest NEP work plan with many recent advances in evaluability assessment.

APPENDIX 3

PROGRAMS VISITED

Location	Program Name	Auspices	Level Of Training	Dates Visited
Wellesley, Massachusetts	Command Training Institute Session #108	New England Institute for Law Enforcement Management/New England Association of Chiefs of Police	Supervisory	January 16-18, 1979
Bay Minette, Alabama	FBI Field Police Training Program, Management Course Level I	Southwest Alabama Police Academy	Mid-Management	January 30 - February 1, 1979
Richmond, Kentucky	Management Level I	Kentucky Bureau of Training	Mid-Management	February 6-8, 1979
Quantico, Virginia	Field Police Training Programs*	Federal Bureau of Investigation (FBI)	Mid-Management and Executive	February 14, 1979
Washington, D.C.	National Sheriff's Institute*	National Sheriff's Association	Executive	February 15, 1979 (November 15, 1978)
Chicago, Illinois	Preservice Superior Officer Training Program	Training Division, Chicago Police Department	Mid-Management	February 22-23, 26, 1979
Evanston, Illinois	Unit Courses*	Northwestern University Traffic Institute	Supervisory	February 26-27, 1979
Washington, D.C.	FBI Police Management and Supervision Course	Washington, D.C. Police Training Academy	Mid-Management	February 23, 26-27; March 1-2, 1979
Monmouth, Oregon	Middle Management Course	Oregon Board of Police Standards and Training	Mid-Management	March 26-29, 1979
Rapid City, South Dakota	Executive Certification Program	South Dakota Criminal Justice Training Center	Executive	March 27-29, 1979
Arlington, Virginia	Police Executive Institute	Police Foundation	Executive	April 1-5, 1979
Winter Park, Florida	Career Development Executive Development Seminar, Break-out Session 1	Rollins College	Executive	April 4-6, 1979
Oakland, California	POST Supervisory Course	Training Division, Oakland Police Department	Supervisory	April 9-12, 1979
Sanford, Florida	Career Development Middle Management Course	Seminole Community College	Mid-Management	April 9-13, 1979
University Park, Pennsylvania	POLEX (Police Executive Development Institute) XLVI	Pennsylvania State University College of Human Development	Executive	May 7-10, 1979
Houston, Texas	Interface 1979	Training Division, Houston Police Department	Supervisory	May 14-16, 1979

* An asterisk indicates that the training process was not observed at a given site. At only three of sixteen sites was the training process not formally observed.

SITE VISIT REPORT OUTLINE

A. Data Management Information

1. Data Sources
2. Classes Observed
3. Investigators on Site
4. Dates on Site
5. Problems in Acquisition or Interpretation of Data

B. Program Profile

1. Full Name of Program
2. Institutional Setting
3. Auspices
4. Training Providers
5. Level of Training
6. Rank of Trainees
7. Jurisdiction of Trainees
8. Incentives for Participation
9. First Offered
10. Persons Trained in Average Session
11. Hours Training in Average Session
12. Times Offered in Average Year
13. Persons Trained in Average Year
14. Hours Training in Average Year
15. Sources of Program Funds
16. Training Expenses met by Participants' Departments
17. Budget Allocation
18. Bases for Program Scheduling
19. Trainee Selection Standards
20. Trainer Credential Requirements
21. Trainer Backgrounds
22. Requirements to Obtain Certificate of Completion
23. Requirements to Obtain Academic Credit, if given
24. Program Aspects Requiring POST Certification
25. Classes Offered
26. Personnel Responsible for Training Coordination
27. Other Levels or Types of Management Training Offered
28. Relationship of Other Levels or Types to the Program
29. Relative Priority Given to the Program
30. Differences in Thrust of Expectations
31. Importance of this Program Relative to Others to the Target Population

C. Origins and History

1. Key Personnel in Design and Implementation
2. Impetus for Program Development
3. Similar Training Programs Accessible at Implementation
4. Model Programs Influencing Program Development
5. Legislative Authority
6. Key Resources Utilized in Implementation
7. Original Developmental Process
8. Major Changes within or Affecting the Program

D. Current Process for Program Development and Management

1. Relationship of Program Development to Management of Departments Served
2. Coordination of Program with others under the same Auspices or Sponsor
3. Central Availability of Information about Training History of Target Pop.
4. Obtaining Broad Input in Development of Program Goals
5. Assessment of Training Needs
6. Setting and Operationalizing Objectives
7. Designing a Program to Serve Objectives
8. Pretesting Trainees
9. Conducting a Program that Meets Objectives and Serves Needs
10. Evaluating In-program Outcomes
11. Evaluating Transfer and Impact
12. Utilization of Program Evaluations
13. Major Anticipated Changes in the Developmental Process

E. Exogenous Factors Affecting Program Development

As Identified

F. Client Flow Models

1. Intended (Testable) Model
   a. Underlying Assumptions
   b. Logical Chain of Assumptions

2. Observed (Equivalency) Model
   a. Underlying Assumptions
   b. Logical Chain of Assumptions
   c. Explication of and Commentary on Observed Model
   d. Exogenous Factors Affecting the Process

3. Differences within the Intended Model
4. Differences within the Observed Model
5. Differences between Intended and Observed Models
6. Implications for Construction of an Evaluable Model

G. Evidence of Effectiveness

1. Measures Used
2. Assessment of Measures Used
3. Available Evaluation Data
4. Assessment of Available Evaluation Data
5. Salient Questions Reflected in Client Flow Models

OMB Number 43S79018

POLICE MANAGEMENT TRAINING PROGRAM SURVEY

NATIONAL EVALUATION PROGRAM PHASE I ASSESSMENT OF POLICE MANAGEMENT TRAINING

The purpose of this survey is to find out how police management training programs are developed, delivered, and evaluated. The aggregated results of this survey will be used in the National Evaluation Program study to help persons operating programs to better manage them. Your individual responses will be held confidential. While you are not required to respond, your cooperation is needed to make the results of this survey comprehensive, accurate, and timely. This report is authorized by law (P.L. 94-503, Sec. 402).

We ask you to complete the survey if your program has been conducted at least once in the past year, and if it fits this broad definition:

An instructional sequence
Under 200 hours in duration
Offered outside a degreed academic program
To currently active or soon to be commissioned police supervisors or managers
Operating at the state or local level
For the purposes of upgrading one or more aspects of supervisory or management performance
For the ultimate benefit of a police agency.

While you may operate more than one program that fits our broad definition, please complete this survey for only the ______________________ course.

Because we are trying to understand current practice in program development, it is important that your answers reflect what you actually do or believe, and not what you think you ought to do or believe. So if a question were to ask, for example, about the extent to which an action is done, and it is not done but you think it should be, tell us that it is not done.

The survey contains 34 questions in total. There are three different types of questions in the survey, so it will be important to observe the directions for each. The three types of instructions ask you to:

Check the one response that is most appropriate

Check all responses that happen to be appropriate

Rate each response on a scale from 0 to 4, to indicate the extent or degree to which the response is accurate. A rating of "0" indicates a response is not at all accurate, "1" connotes low accuracy, "2" moderate accuracy, "3" high accuracy, and "4" total or nearly complete accuracy. When two or more alternative responses are equally accurate, it is correct to give them equal ratings. The exact focus of the ratings process will vary from question to question, but will in each case be made explicit.

If you have any other comments related to the survey, we are interested in hearing them. Please place these comments on additional sheets and enclose them with the completed survey. In addition, if your program has completed any follow-up evaluations of training's effects on behavior or training's worth to the organization, please enclose these as well.

The notations in the right margin are for later office use and may be ignored.

If you have not already done so, please complete and send back the postage-paid return post-card enclosed before filling out the survey.

BEGIN CARD [0][1] 1-2
#[ ][ ][ ] 3-5

1. For what level target audience is the program primarily intended? (CHECK ONE; if more than one program is offered, you should fill out the survey for only the one listed on the first page) 6-10

[ ] a. First line supervisors
[ ] b. Middle managers
[ ] c. Executives
[ ] d. A cross section of the three above
[ ] e. Other (specify) ____________________

2. How many times in the past year has the program been conducted? (CHECK ONE) 11-17

[ ] a. Not at all (If you checked this box, do not complete rest of survey)
[ ] b. Once
[ ] c. Twice
[ ] d. Three times
[ ] e. Four to six times
[ ] f. Seven to ten times
[ ] g. More than ten times

3. How many trainees have taken part in the average program session over the past year? (CHECK ONE) 18-24

[ ] a. 1-5
[ ] b. 6-10
[ ] c. 10-15
[ ] d. 16-20
[ ] e. 21-30
[ ] f. 31-40
[ ] g. More than 40

4. From what type of jurisdiction does the program draw its primary audience? (CHECK ONE) 25-31

[ ] a. Single municipality
[ ] b. Single county
[ ] c. Single metropolitan area
[ ] d. Intrastate region
[ ] e. Statewide
[ ] f. Multistate
[ ] g. National

CONTINUE CARD 01

5. Under what auspices is the program offered? (CHECK ALL APPROPRIATE) 32-41

[ ] a. State Planning Agency (SPA)
[ ] b. Police standards and training agency in the state
[ ] c. Academic institution
[ ] d. Municipal or county academy
[ ] e. Regional academy
[ ] f. State academy
[ ] g. FBI
[ ] h. State associations
[ ] i. National associations
[ ] j. Other (specify) ____________________

6. To what extent are funds to support the program drawn from these sources? (RATE EACH on a scale from 0 through 4, with 0 indicating funds were not drawn from a source, 4 that funds were drawn totally or almost totally from a source, and the points in between intermediate levels of dependence for funds) 42-49

[ ] a. Local departmental budget
[ ] b. Tuition and fees
[ ] c. State allocation for police training
[ ] d. State penalty assessment fund for police training
[ ] e. State Department of Education reimbursement for vocational training
[ ] f. Federal funds directly from LEAA
[ ] g. LEAA funds channeled through the SPA
[ ] h. Federal funds from the FBI, DOT, or another agency other than LEAA

7. Which aspects of the program, if any, must be certified by the state's police standards and training agency? (CHECK ALL APPROPRIATE) 50-57

[ ] a. Objectives
[ ] b. Curricula
[ ] c. Trainers
[ ] d. Location
[ ] e. Date
[ ] f. Entire program as a unit
[ ] g. Other (specify) ____________________
[ ] h. None

8. To what extent did these individuals and groups have an influence in establishing the broad purposes, or goals, of the program? (RATE EACH on a scale from 0 through 4, with 0 indicating no influence, 4 total or near total influence, and the points in between intermediate levels of influence) 58-73

[ ] a. Program staff
[ ] b. Program instructors
[ ] c. Graduates of related programs
[ ] d. Program advisory board
[ ] e. Police chiefs and executives from user agencies
[ ] f. Training directors from user agencies
[ ] g. SPA

CONTINUE CARD 01

[ ] h. Police standards and training agency in the state
[ ] i. State legislature
[ ] j. Local government officials
[ ] k. National police associations
[ ] l. Police unions
[ ] m. State police associations
[ ] n. Citizens groups and public interest groups
[ ] o. Contracted outside organizations
[ ] p. Local academic institutions

9. What techniques have been used to clarify the nature of the performance deficiencies giving rise to the need for training? (CHECK ALL APPROPRIATE) 74-80

BEGIN CARD [0][2] 1-
REPRO 3-5

[ ] a. Informal interviews with chiefs and other user group executives 6-
[ ] b. Informal interviews with program graduates
[ ] c. Posttraining reaction surveys of trainees from prior sessions
[ ] d. Formal surveys of target audience incumbents
[ ] e. Formal surveys of chiefs and user group executives
[ ] f. Formal surveys of entire agencies
[ ] g. Formal delphic surveys of experts
[ ] h. Job-task analysis
[ ] i. Individual testing of target audience
[ ] j. Management audits of participating agencies
[ ] k. Other (specify) ______________________

10. To what extent were these already existing information sources used to determine training needs? (RATE EACH on a scale from 0 through 4, with 0 indicating that a source was not used at all, 4 indicating that a source was used in totality, and the points in between intermediate levels of use) 10-17

[ ] a. Centralized data on the training histories of the target population
[ ] b. Course specifications and curricula outlines from recognized programs
[ ] c. National standards on police manager needs, including the Police Executive Report
[ ] d. National Manpower Survey (NMS)
[ ] e. Other national or state studies of police manager effectiveness
[ ] f. Job-task analyses borrowed from other jurisdictions
[ ] g. Other forms of needs assessments borrowed from other jurisdictions
[ ] h. Needs assessments developed for prior programs in this jurisdiction

11. Which of these means are used to help guide the instructional process? (CHECK ALL APPROPRIATE) 18-

[ ] a. General outline of curricula topics
[ ] b. Module-by-module course content summaries
[ ] c. Broadly articulated course objectives
[ ] d. Objectives showing in measurable terms the areas in which change is desired and the intended extent of change (terminal performance objectives)
[ ] e. Other (specify) ______________________

CONTINUE CARD 02

12. Which of these factors, if any, have reduced or even eliminated the need for a more formal process of goal formation, needs assessment, and objective setting? (CHECK ALL APPROPRIATE) 23-28

[ ] a. Close relationship and understanding between program staff and user groups on an informal basis
[ ] b. Program staff's beliefs that certain needs must be addressed, regardless of whether user groups are aware of these needs
[ ] c. Command orientation of the organization in which the program exists
[ ] d. Organizational training priorities
[ ] e. Legal requirements and other mandates imposed on the program
[ ] f. Other (specify) ______________________

13. To what extent are these criteria important in trainee selection? (RATE EACH on a scale from 0 through 4, with 0 indicating no importance, 4 total and predominant importance, and the points in between intermediate levels of importance) 29-38

[ ] a. Scores on promotional exams
[ ] b. Time in grade
[ ] c. Rank
[ ] d. Current responsibilities
[ ] e. Prior education
[ ] f. Time elapsed since last participation in a career sequence
[ ] g. Other prior training
[ ] h. Recommendation by supervisors
[ ] i. "Promise" for higher levels of responsibility
[ ] j. Individual's demonstrated training needs

14. To what extent are these criteria important in trainer selection? (RATE EACH on a scale from 0 through 4, with 0 indicating no importance, 4 total and predominant importance, and the points in between intermediate levels of importance) 39-51

[ ] a. Advanced academic achievement
[ ] b. Law enforcement experience
[ ] c. Immediate rank in department
[ ] d. Experience directly related to subject area
[ ] e. Completion of an instructor's course
[ ] f. Submission of a sample lesson plan, including performance objectives
[ ] g. University teaching experience
[ ] h. Experience as a trainer in the subject area
[ ] i. Sincere interest in teaching police officers
[ ] j. National recognition on the subject to be presented
[ ] k. Graduation from the program
[ ] l. Congruence with the program's philosophy, based on oral interview
[ ] m. Personal recommendation by another instructor

CONTINUE CARD 02

15. What types of advance information are collected about trainees or their departments before or at the start of training? (CHECK ALL APPROPRIATE) 52-65

[ ] a. Biographical information covering such items as law enforcement experience, educational achievement, prior related training, etc.
[ ] b. Trainee's description of his key management problems
[ ] c. Trainee's personal and career goals
[ ] d. Pretraining skill or knowledge levels
[ ] e. Pretraining attitudinal or personality measures
[ ] f. Pretraining measures of trainee's on-the-job behavior
[ ] g. Trainee perceptions of their training needs and/or management style
[ ] h. Superior ratings of trainee's training needs and/or management style
[ ] i. Peer ratings of trainee's training needs and/or management style
[ ] j. Subordinate ratings of trainee's training needs and/or management style
[ ] k. Trainee's self-selected objectives for what to gain in training
[ ] l. Objectives set for trainees by superiors as to what to gain in training
[ ] m. Pretraining measures on key indicators of overall departmental performance
[ ] n. Other (specify) ______________________

16. To what extent is advance information shown in the question above used in the following ways? (RATE EACH on a scale from 0 through 4, with 0 indicating non-use, 4 indicating near total or complete use, and the points in between intermediate levels of use) 66-

[ ] a. Select trainees and assign them to appropriate sessions
[ ] b. Provide a comparative basis for later program evaluations
[ ] c. Adjust the overall level of presentation
[ ] d. Add supplementary exercises and materials to focus on the overall immediate needs of a class
[ ] e. Give individual attention to trainees
[ ] f. Provide a framework for trainee development of personal Action Plans

17. Through what means is the advance information described above collected? (CHECK ALL APPROPRIATE) 72-

[ ] a. Applications
[ ] b. Formal pretraining surveys
[ ] c. Initial meeting of trainees with program staff
[ ] d. Assigned prework
[ ] e. Individualized testing
[ ] f. Direct observation of performance on the job
[ ] g. Review of departmental performance data
[ ] h. Other (specify) ______________________

BEGIN CARD [0][3] 1-2
REPRO 3-5

18. To what extent are the following means used to coordinate trainer activities and ensure compliance with program curricula? (RATE EACH on a scale from 0 through 4, with 0 indicating non-use, 4 indicating near total or complete use, and the points in between intermediate levels of use) 6-10

[ ] a. Observation of each new instructor in class
[ ] b. Periodic spot-observation of all instructors in class
[ ] c. Periodic faculty meetings
[ ] d. Periodic interviews by the program director with each faculty member
[ ] e. Periodic review of lesson plans

19. What minimum requirements must a trainee fulfill merely to obtain a certificate of program completion? (CHECK ALL APPROPRIATE) 11-21

[ ] a. Attend a specified percentage of classes
[ ] b. Participate actively in all attended sessions
[ ] c. Take one or more exams
[ ] d. Pass one or more exams [if you check d, do not check c]
[ ] e. Write a paper
[ ] f. Submit a notebook (record of class notes)
[ ] g. Make a class presentation
[ ] h. Display certain desired behaviors to the satisfaction of the instructor
[ ] i. Complete a group project
[ ] j. Provide evidence a few weeks after the program of implementation efforts
[ ] k. Other (specify) ______________________

20. Over and above the minimum requirements to obtain a certificate of course completion, what additional requirements must a trainee fulfill to obtain academic credit, if this option exists? (CHECK ALL APPROPRIATE) 22-33

[ ] a. Attend a specified percentage of classes
[ ] b. Participate actively in all attended sessions
[ ] c. Take one or more exams
[ ] d. Pass one or more exams [if you check d, do not check c]
[ ] e. Write a paper
[ ] f. Submit a notebook (record of class notes)
[ ] g. Make a class presentation
[ ] h. Display certain desired behaviors to the satisfaction of the instructor
[ ] i. Complete a group project
[ ] j. Provide evidence a few weeks after the program of implementation efforts
[ ] k. Pay an additional fee
[ ] l. Other (specify) ______________________

21. During or immediately at the program's completion, what methods are used to evaluate the program? (CHECK ALL APPROPRIATE) 34-40

[ ] a. Trainee course critiques or reaction surveys
[ ] b. Written examination to test changes in knowledge or skill
[ ] c. Structured observations by instructors of changed trainee behavior in simulation exercises

CONTINUE CARD 03

[ ] d. Development of an Action Plan based on program contents
[ ] e. Measures of attitudinal change on a paper-and-pencil personality test
[ ] f. Assessment of trainee performance by other trainees
[ ] g. Other (specify) ______________________

22. To what extent have the evaluations performed during the program or immediately after its completion been used in the following ways? (RATE EACH on a scale from 0 through 4, with 0 indicating non-use, 4 indicating near total or complete use, and the points in between intermediate levels of use) 41-

[ ] a. Alter logistics (e.g., housing arrangements, meals, class seating)
[ ] b. Eliminate unpopular courses and change topical emphases
[ ] c. Alter instructional techniques
[ ] d. Modify the order or sequencing of course modules
[ ] e. Expand or decrease the use of particular instructional personnel
[ ] f. Revise and update the course description
[ ] g. Change the name of the program
[ ] h. Change trainee selection standards
[ ] i. Change trainer selection standards
[ ] j. Implement or refine auditing procedures
[ ] k. Refine the process of assessing needs and setting objectives
[ ] l. Modify goals and objectives
[ ] m. Increase emphasis on transfer and implementation techniques
[ ] n. Change the nature or extent of contact with user groups prior to training
[ ] o. Reduce the inhibiting effects of forces external to the program on program development and delivery
[ ] p. Develop or expand management training options at other levels of management
[ ] q. Start educational efforts to reconcile discrepant views of training needs

23. After the trainee has returned to the job and had a reasonable chance to use what he learned, what types of follow-up evaluations of behavior or results do you conduct? (CHECK ALL APPROPRIATE) 58-74

[ ] a. Trainee's self-assessment of training utilization
[ ] b. Superior, subordinate, or peer assessments of training utilization
[ ] c. Trainee's reassessment of his training needs
[ ] d. Superior, subordinate, or peer reassessment of trainee's training needs
[ ] e. Trainee's self-assessment of changes in management style
[ ] f. Superior, subordinate, or peer assessment of changes in management style
[ ] g. Trainee's self-assessment of progress in implementing the Action Plan
[ ] h. Superior, subordinate, or peer assessments of trainee's progress in implementing the Action Plan
[ ] i. Measures of attitudinal change on a paper-and-pencil personality test
[ ] j. Retesting to determine retention of course contents
[ ] k. Structured observation of simulated job situation

CONTINUE CARD 03

[ ] l. Structured observation of actual on-the-job behavior
[ ] m. Paper-and-pencil assessments of changes in organizational functioning
[ ] n. Organizational management audits
[ ] o. Measures of overall organizational productivity
[ ] p. Cost-benefit analyses of training's worth to the organization
[ ] q. Other (specify) ______________________

24. To what extent have follow-up evaluations been used in these ways? (RATE EACH on a scale from 0 through 4, with 0 indicating non-use, 4 indicating near total or complete use, and the points in between intermediate levels of use) 75-80

BEGIN CARD [0][4] 1-2
REPRO 3-5

[ ] a. Alter logistics (e.g., housing arrangements, meals, class seating) 6-16
[ ] b. Eliminate unpopular courses and change topical emphases
[ ] c. Alter instructional techniques
[ ] d. Modify the order or sequencing or course modules
[ ] e. Expand or decrease the use of particular instructional personnel
[ ] f. Revise and update the course description
[ ] g. Change the name of the program
[ ] h. Change trainee selection standards
[ ] i. Change trainer selection standards
[ ] j. Implement or refine auditing procedures
[ ] k. Refine the process of assessing needs and setting objectives
[ ] l. Modify goals and objectives
[ ] m. Increase emphasis in training on transfer and implementation techniques
[ ] n. Change the nature or extent of contact with user groups prior to training
[ ] o. Reduce the inhibiting effects of forces external to the program on program development and delivery
[ ] p. Develop or expand management training options at other levels of management
[ ] q. Start educational efforts to reconcile discrepant views of training needs

25. To what extent have these factors negatively affected the adequacy of program funding? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating near total and predominant effect, and the points in between intermediate levels of effect) 17-23

[ ] a. Reduction in state allocation to training
[ ] b. Imposition of a cap on expenditures
[ ] c. Reduction in the availability of trainees able to pay tuition
[ ] d. Requirement that all mandated programs must be conducted before residual funds may be expended for other purposes
[ ] e. Definitional problems affecting use of in-service funds for management training
[ ] f. Lack of a line item in the budget for training
[ ] g. Dependence on multiple agencies for financial support

CONTINUE CARD 04

26. To what extent have funding restrictions affected program development and delivery at each of these points? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating near total and predominant effect, and the points in between intermediate levels of effect) 24-27

[ ] a. Goal setting and needs assessment
[ ] b. Objective setting and curriculum design
[ ] c. Delivery
[ ] d. Evaluation

27. If the program must operate within legal requirements imposed by the state's police standards and training agency, to what extent do the following statements describe the requirements? (RATE EACH on a scale from 0 through 4, with 0 indicating complete inaccuracy, 4 indicating complete accuracy, and the points in between intermediate levels of accuracy) 28-31

[ ] a. They were developed by following a closely reasoned process or standard
[ ] b. They are set forth by the oversight agency as the minimum basis for program development, but not as the full scope of developmental activities
[ ] c. They are regularly reviewed on a formal basis to assess their current applicability
[ ] d. They include a mechanism for their enforcement

28. To what extent have unions affected the program in these ways? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 near total and predominant effect, and the points in between intermediate levels of accuracy) 32-

[ ] a. Promulgated conditions under which training may be offered
[ ] b. Restructured selection procedures so no one may be discriminated against on the basis of ability
[ ] c. Influenced course content and topical emphases
[ ] d. Controlled attendance by defining the training day as in excess of eight hours and hence justifying overtime pay
[ ] e. Interfered with follow-up evaluations by telling trainees their program evaluations would be used personally as performance appraisals

29. If the program is based in an academic institution, has this host institution imposed any requirements upon the program? (CHECK ALL APPROPRIATE) 37-40

[ ] a. Modify curricula
[ ] b. Modify entrance requirements and procedures
[ ] c. Modify testing requirements
[ ] d. Other (specify) ________________________________

CONTINUE CARD 04

30. In developing the program, to what extent were resources from existing programs used, instead of developing and tailoring resources specifically for the program? (RATE EACH on a scale from 0 through 4, with 0 indicating non-use, 4 indicating near total and complete use, and the points in between intermediate levels of use) 41-48

[ ] a. Goal statements
[ ] b. Needs assessments
[ ] c. Performance objectives
[ ] d. Curricula
[ ] e. Success measures
[ ] f. Facilities
[ ] g. Trainers
[ ] h. Evaluation strategies

31. To what extent have the following forces affected the program's ability to perform a more formal type of needs assessment? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating near total and predominant effect, and the points in between intermediate levels of effect) 49-54

[ ] a. Lack of skilled personnel
[ ] b. Lack of lead time to prepare the program step-by-step
[ ] c. Lack of computer facilities
[ ] d. Lack of funds
[ ] e. Lack of research techniques and designs
[ ] f. Lack of measures and concepts for analyzing performance

32. To what extent have these factors affected availability of trainers who fit the program's selection standards? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating near total and predominant effect, and the points in between intermedidate levels of effect) 55-58

[ ] a. Lack of qualified instructors in the jurisdiction
[ ] b. Lack of lead time to obtain appropriate instructional services
[ ] c. Political pressures on selection
[ ] d. Institutional pressures on selection

33. To what extent have these factors restricted the availability of trainees for which the program was designed? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating near total and predominant effect, and the points in between intermediate levels of effect) 59-66

[ ] a. Coverage requirements in the agencies served by the program
[ ] b. Tuition fund availability
[ ] c. Union demands for overtime pay during training
[ ] d. Pressures to make greater use of local programs
[ ] e. Adequacy of incentive funds to attract participants
[ ] f. Relative priority of training in participating agencies
[ ] g. Saturation of the local audience
[ ] h. Pressures on the program to maximize enrollments to justify its budget

CONTINUE CARD 04

34. To what extent do the following factors limit the amount and quality of program evaluations? (RATE EACH on a scale from 0 through 4, with 0 indicating no effect, 4 indicating total and near predominant effect, and the points in between intermediate levels of effect) 67-73

[ ] a. The program's expectations about what it intends to accomplish are not clearly defined
[ ] b. The program's expectations or objectives are not really plausible, because resources are inadequate or ineffectively used
[ ] c. There is little or no agreement over what variables ought to be measured
[ ] d. Existing data systems are not generally reliable
[ ] e. Evaluation costs are out of line with their potential uses
[ ] f. The political burdens or threats posed by evaluations make them unfeasible
[ ] g. Evaluations are not likely to have much effect on program operations

PROGRAM PERFORMANCE MEASURES

What types of measurements could be taken to evaluate program performance? When conditions warrant evaluation, techniques and measures exist to construct an evaluation strategy offering reasonable levels of rigor in design and confidence in results. Most of the outcome and impact variables in Chapter Four's models have not been measured in police management training programs. In fact, many of the variables have rarely been measured even in the industrial management training programs where training evaluation techniques have largely developed. This results from dissociation of the training function from the evaluation of training's effect. The purposes of this appendix are to outline the key measurement points synthesized from Chapter Four's models and to identify potential approaches to measurement and data collection.

This appendix presents program performance measures in six tabular exhibits, each divided into four sections. The six exhibits group performance measures in terms of knowledge and skill outcomes (Exhibit 6-A), attitudinal outcomes and impacts (Exhibit 6-B), individual behavioral impacts (Exhibit 6-C), career development outcomes and impacts (Exhibit 6-D), departmental and systemic impacts (Exhibit 6-E), and program-related outcomes and impacts (Exhibit 6-F). The four sections in each exhibit contain: the key data collection and analysis techniques for a given group of variables, selected model referents germane to the grouping, partial indicators of success corresponding to the listed model referents, and alternative measurements that target the generic measures to the partial indicators.

- <u>Data Collection And Analysis Techniques</u>--Includes several broad measures adaptable to the needs and resources of a particular evaluation.

- <u>Model Referents</u>--Draws selectively upon comprehensive tables of variables found in the models to display those of widest interest.

- <u>Partial Indicators</u>--Contains general translations of the model referents into quantitative success indicators. These are still only <u>partial</u> indicators because the appropriate time-frames and acceptable range of values for performance can be determined only by a specific program.

- <u>Alternate Measurements</u>--Shows which of the generic measures are most relevant to the indicator and brings the indicator closer to a measurable state, in some instances by suggesting specific instrumentation.

We intentionally avoid repeated reference in the exhibits to the data with which post-training performance should be compared because most measures can appropriately use three comparative bases:

EXHIBIT G-A(1)

NEP/Police Management Training

KNOWLEDGE AND SKILL OUTCOMES

Key Data Collection and Analysis Techniques for Knowledge and Skill Outcomes

(1) Content analyses of written and oral course critiques for understanding of course content

(2) Structured observation and content analysis of class discussions

(3) Paper-and-pencil exams based on content and requiring desired type and level of mastery

(4) Content analysis of case studies calling for use of information and principles from training

(5) Structured observation of skill demonstrations

(6) Structured observation of simulation exercises and structured group experiences

(7) Content analysis of written exercises requiring the simulation of on-the-job behavior

(8) Content analysis of Action Plans developed by trainees at the conclusion of training

(9) Scores on self-assessment inventories

Selected Model Referents	Partial Indicators	Alternate Measurements
Knowledge and understanding of minimum required duties and responsibilities specific to rank, and of the legal, administrative, and procedural scope of duties	# and % of trainees demonstrating during or at the completion of training increased knowledge of their minimum required duties specific to rank, and the general legal, administrative and procedural scope of duties; the extent of increase	2;3 (preferably based on departmental task analysis)
Knowledge and understanding of the external changes in laws, regulations, and policies necessitating adaptation; of current practices that are unacceptable; of recommended practices to adapt to change	# and % of trainees demonstrating during or at the completion of training increased knowledge of requirements imposed by changes in the legal network, the current practices unacceptable in light of these changes, and the practice recommended to adapt to change; the extent of increase	2;3;4;6;7;8
Knowledge of business management theory and practice; or of participative management theories, their interrelationships, and related practice	# and % of trainees demonstrating during or at the completion of training increased knowledge of general or participative management theory and practice; the extent of increase	1;2;3;4;6;7;8
Knowledge of the advantages and disadvantages of alternative practices in comparison with current practice	# and % of trainees demonstrating during or at the completion of training increased knowledge of the advantages and disadvantages of a proposed set of alternative practices as compared with current practice; the extent of increase	1;2;3

EXHIBIT 6-B

NEP/Police Management Training

ATTITUDINAL OUTCOMES AND IMPACTS

Key Data Collection and Analysis Techniques for Attitudinal Outcomes and Impacts		
(10) Scores on general personality inventories	(11) Scores on instruments measuring particular components of management behavior	
Selected Model Referents	**Partial Indicators**	**Alternate Measurements**
Assimilation and internalization of experience with alternate management systems	# and % of trainees demonstrating at the completion of training, or after a specified time back on the job, changed attitudes in the direction of those espoused by alternate management systems; extent of change	10;11 (using the Management Grid, Personnel Relations Survey, Leadership Opinion Questionnaire, Leadership Description Behavior Questionnaire, Responsibility Authority Delegation Scale and/or Supervisory Behavior Description, etc.)
Reduced training insularity	# and % of trainees demonstrating at the completion of training, or after a specified time back on the job, a reduction in insularity from other managers within and outside the department, and from the general public; the extent of change	10;11

Selected Model Referents	Partial Indicators	Alternative Measurements
Knowledge of the preconditions for, problems in, and precedents for successful implementation of alternative practices	# and % of trainees demonstrating during or at the completion of training increased knowledge of the preconditions for, problems in, and precedents for successful implementation of alternative practices; the extent of increase	1;2;3;6;7;8
Knowledge of a prospective decision and the problem it addresses	# and % of trainees demonstrating during or at the completion of training increased knowledge of a prospective decision and the problem it addresses; the extent of increase	1;2;3;6;8
Self-knowledge on key dimensions of management practice	# and % of trainees demonstrating during or at the completion of training greater familiarity with their own management styles and training needs; the extent of increased familiarity; especially, posttraining self-assessments more in line than pretraining self-assessments with pretraining assessments by co-workers.	2;6;9
Ability to develop clear and measureable objectives	# and % of trainees demonstrating during or at the completion of training increased ability to set clear and measureable objectives; the extent of increase	5;6;8
Capacity to analyze hypothetical situations in terms of participative as well as more traditional management principles	# and % of trainees demonstrating during or at the completion of training increased ability to analyze situations in terms of several management theories; the extent of increase	2;4;6;7
Capacity for informed critique of a prospective decision	# and % of trainees demonstrating during or at the completion of training increased capacity to assess a prospective decision in terms of departmental performance needs, results of prior attempts at implementation, and relationship of decision to current philosophy and practice; the extent of increase	1;2;8
Ability to recognize and analyze situations requiring adaptation	# and % of trainees demonstrating during or at the completion of training increased ability to recognize and analyze situations requiring adaptation	2;4;7;8

[illegible]IT 6-[illegible]

Selected Model Referents	Partial Indicators	Alternative Measurements
Obtaining information and assistance from other program graduates and program staff in management problem solving and decision making	# and % of trainees reporting receipt or provision of information or assistance from or to other program graduates in postprogram problem solving and decision-making; # and % of trainees receiving similar information or assistance from program staff; extent of this information and assistance; especially, greater use of nondepartmental sources of assistance than by comparable nontrainees	17;18;24
Limited implementation of specific management practices in own domain	# and % of trainees demonstrating implementation of specified management practices after a specified posttraining interval has elapsed; the extent of implementation	13;20;22 (using the Management Grid or Johari Window to measure participation practices)
Refinement of individual management style	# and % of trainees demonstrating changes in an intended specified direction in stated aspects of management style; the extent of change	13;14;22 (using the Management Grid, Personnel Relations Survey, Leadership Opinion Questionnaire, Leadership Description Behavior Questionnaire, Responsibility Authority Delegation Scale, and/or Supervisory Behavior Description, etc.)
Hightened individual effectiveness	# and % of trainees demonstrating effectiveness in their agency in intended specified ways; the extent of change	13;14;16;20;21;23

NEP/Police Management Training

INDIVIDUAL JOB PERFORMANCE IMPACTS

Key Data Collection and Analysis Techniques for Individual Job Performance Impacts

(12) Follow-up retest on retention (and presumed use) of program contents at a desired level of mastery

(13) Follow-up structured observation of simulated or actual job performance

(14) Follow-up content analysis of case studies calling for use of information and principles from training

(15) Standard departmental compliance measures (e.g., complaints)

(16) Standard departmental morale measures (e.g., absenteeism, grievances filed, lateness, turnover)

(17) Content analysis of diaries, logs, duty reports, or correspondence

(18) Content analysis of program records and correspondence

(19) Content analysis of information used in Action Plan follow-up

(20) Follow-up surveys on actual implementation of Action Plan--including those formulated through mutual trainee-superior agreement--to trainees, superiors, peers and/or subordinates

(21) Surveys of incumbents and/or supervisors about the compliance of those completing training with the minimal duties and responsibilities of their positions

(22) Surveys of trainees, supervisors, peers and/or subordinates about changes in specific aspects of individual management behavior or style

(23) Surveys of posttraining needs for training using the same standardized instrument or an agency-specific task analysis administered to trainees, superiors, peers and/or subordinates before training (e.g., managerial training needs profile)

(24) Surveys of trainees about postprogram information or assistance received from or provided to other program graduates

Selected Model Referents	Partial Indicators	Alternate Measurements
Retention and use of program knowledge in on-the-job discussions	# and % of trainees retaining knowledge of program contents and using program language on the job after a specified posttraining interval has elapsed; extent of retention and use	12;14;17;19
Short-term performance boost	# and % of trainees displaying greater overall output during a specified interval immediately after return from training; extent of increased output	16;21
Performance of new duties in compliance with departmental regulations and policies; renewed performance of duties in compliance with regulations and policies	# and % of trainees displaying increased compliance in performance of their new or continuing duties with departmental regulations and policies; the extent of compliance; especially in the case of those newly promoted, greater compliance than a comparable nontrained population	13;15;21

EXHIBIT 6-E(1)

NEP/Police Management Training

DEPARTMENTAL AND SYSTEMIC IMPACTS

Key Data Collection and Analysis Techniques for Departmental and Systemic Impacts

(32) Standard measures of morale (e.g., absenteeism, grievances filed, lateness, turnover)

(33) Standard measures of compliance (e.g., complaints)

(34) Surveys of departmental personnel about morale

(35) Surveys of departmental personnel about compliance with departmental regulations and policies

(36) Surveys of departmental personnel about adaptation to Federal and state laws and regulations

(37) Surveys of departmental personnel about specific department-wide changes in management practice

(38) Surveys of departmental personnel about reduction in specific management problems

(39) Surveys of departmental personnel about assistance given by graduates to other graduates in implementing organizational change

(40) Surveys of departmental personnel about feasibility of and need for organizational change

(41) Surveys of departmental personnel about posttraining department-wide needs for training, using the same standardized instrument administered prior to training (e.g., Managerial Training Needs Profile)

(42) Follow-up surveys on implementation of Action Plan department-wide

(43) Analysis of case records for court challenges to departmental activities

(44) Departmental management audits

(45) Surveys of other state or local agencies about compliance with regulations or about agency effectiveness

(46) Citizen surveys about the quality of police services

(47) Rates at which trainees fail in the program

(48) Follow-up on career development of program failures

(49) Analysis of background data from entry level and laterally-transferring police personnel

(50) Departmental or system turnover rates

Selected Model Referents	Partial Indicators	Alternate Measurements
General morale boost	# and % of departmental personnel experiencing an overall increase in morale; extent of increase in morale	32;34
Departmental control	# and % of departmental personnel acting in general compliance with agency regulations and policies; the extent of compliance	33;35;44;46
Departmental adaptation to external changes in the net of legal requirements	# and % of departmental personnel acting in general compliance with external changes in the net of legal requirements; extent of compliance	36;43;44;45;46

EXHIBIT 6-D

NEP/Police Management Training

CAREER DEVELOPMENT OUTCOMES AND IMPACTS

Key Data Collection and Analysis Techniques for Career Development Outcomes and Impacts		
(25) Analysis of the career records of program graduates for promotions, duty assignments, training and educational histories, other professional activities (26) Follow-up surveys of program graduates about career progression, educational and training activities, reading habits, time use, conference attendance, and other professional activities (27) Follow-up surveys of superiors, peers and subordinates about perceived status of training graduates	(28) Follow-up survey of program staff and program graduates about assistance and support provided to other graduates in career development (29) Analysis of program records and correspondence for evidence of assistance provided to graduates in career development (30) Content analysis of end-of-program Action Plans to detect interest in career development (31) Follow-up surveys about implementation of Action Plans related to career development	
Selected Model Referents	**Partial Indicators**	**Alternate Measurements**
Accelerated career development	# and % of trainees receiving promotions at a faster rate than incumbents not trained by a particular program; # and % attaining higher terminal positions than incumbents not trained by a particular program	25;26;30;31
Participation in career development activities	# and % of trainees taking part in posttraining career development activities, such as additional training and education, conferences, professional reading, etc.; extent of participation	25;26;30;31
Recognition of changed trainee status subsequent to training	# and % of trainees being viewed as "select" by superiors, peers and subordinates immediately after return from training; extent of being viewed as select; # and % of trainees assigned expanded duties immediately after return from training; extent to which duties are expanded	25;27
Maintained acquaintances with program graduates	# and % of trainees maintaining acquaintances with other program graduates a specified duration after completion of training; extent of contact against an absolute standard	28
Support by the program network of the promotion of its graduates	# and % of trainees receiving or providing support from or to other program graduates in promotions; # and % of trainees receiving similar support from program staff; the extent of support	28;29

EXHIBIT 6-F

NEP/Police Management Training

PROGRAM-RELATED OUTCOMES AND IMPACTS

Key Data Collection and Analysis Techniques for Program-Related Outcomes and Impacts

(51) Analysis of program budgets

(52) Records of program operations and services

(53) Records of program attendance and participation in network activities

(54) Surveys of program reputation among law enforcement agencies and graduates

(55) Surveys of network participation by program graduates

Selected Model Referents	Partial Indicators	Alternate Measurements
Program maintenance, growth and development	Extent of increases and projected continuity of program support	51;52
Initiation of trainees into and expansion of the program network	# of trainees added to alumni; # and % of trainees active in alumni activities	53;55
"Advertising" by graduates and subsequent attendance by co-workers at program offerings	# and % of graduates encouraging co-workers to attend; # and % of new trainees indicating a co-worker as the initial trigger for interest in the program	53;54;55
Attendance by program graduates at subsequent program offerings	# and % of trainees having attended prior program offerings; extent of prior participation	53;55
Participation of program graduates in network activities	# and % of graduates taking part in graduate activities, including attendance at reunions and subscription to newsletters; extent of graduate participation	53;55

Selected Model Referents	Partial Indicators	Alternative Measurements
Departmental acceptance of the need for and feasibility of organizational change	# and % of departmental personnel expressing support for organizational change; extent of expressed support	40 (perhaps using Profile of Organizational Characteristics)
Support from program graduates and staff to secure executive commitment for and reduce obstacles to organizational change	# and % of trainees receiving assistance from program staff or graduates, or providing assistance to program graduates, in securing executive commitment for and reducing obstacles to organizational change; extent of assistance	39
Successful implementation of large-scale organizational change	# and % of departmental personnel perceiving that the agency has successfully implemented intended and needed organizational change; extent of such perceived change, broken down by rank	37 (Profile of Organizational Characteristics); 38 (Blockage Questionnaire); 42; 44
Staff stability	# and % of departmental and system personnel increased job tenure, and expressing an intent to remain in their jobs	50
Weeding out of marginal and substandard managers	# and % of trainees failing to complete a required program and/or certification requirements, and either retiring or leaving law enforcement	47;48
Attraction of more qualified personnel to law enforcement	# and % of new departmental or system personnel possessing more credentials of a specified type than their predecessors	49
Statewide law enforcement reputation	# and % of law enforcement agencies, municipal government agencies, and the general public perceiving improvements in the quality of police management; extent of perceived improvements	45;46
Departmental effectiveness	# and % of intended changes in management practice, reduction in management problems, reduction in management training needs, reduction in need for organizational change; extent of changes	37 (Profile of Organizational Characteristics); 38 (Blockage Questionnaire); 40; 41;45;46

GUIDELINES FOR WRITING ACTION ITEMS

A. WHAT AN ACTION ITEM LOOKS LIKE

1. The most important characteristic of an action item is that it is written so that you--or someone else--will <u>know when it occurs</u>. One way to help achieve this is to use specific <u>action verbs</u>.

   The following is a list of such verbs:*

Mental Skill		Physical Skill	Attitude
State	Demonstrate	Execute	Choose
Name	Discriminate	Operate	Volunteer
Describe	Classify	Repair	Allow
Relate	Generate (a solution)	Adjust	Recommend
Tell	Apply (a rule)	Manipulate	Defend
Write	Solve	Handle	Endorse
Express	Derive	Manufacture	Cooperate
Recount	Prove	Calibrate	Accept
	Analyze	Remove	Decide to
	Evaluate	Replace	Agree

2. As you are working on the action items, ask yourself: <u>Is the behavior described observable</u>? Will it be obvious to me or others when it happens?

3. Examples of action items:

   <u>As a result of being in this course, I plan to</u>:

   a. <u>Describe</u> this course to my supervisor within a week of my returning to the job. As a result, my supervisor will know: the contents of the course, how I can apply what I learned to the job, and whether or not others in the organization should attend.

   b. <u>Handle</u> every piece of paper only once in order to improve the management of my own time. Begin as soon as I am back on the job.

   c. <u>Apply</u> the principles of performance analysis to the problem of incomplete and tardy case reviews in my division; request assistance from the training office, as needed. As a result, I will know whether training is required and/or some other solution is appropriate to reduce the problem. Begin within a month upon returning.

   d. <u>Talk</u> with my employees directly about a problem that arises rather than avoiding a confrontation; discuss the situation in order to reach mutual understanding.

   e. Within two weeks after I return, <u>negotiate</u> with my supervisor to implement a ____________ system in my unit and come to an agreement on whether or not I can proceed.

B. IMPLEMENTING THE ACTION ITEM

1. As you proceed to develop action items, be sure to <u>think of yourself in your actual job setting</u>, implementing the activity you have described.

2. If you have an idea of <u>when</u> you will be able to begin implementing the action items, you can make a note of it. Three categories can be chosen: (1) "as arises" (you don't know when the opportunity to try this item will occur), (2) "within 2 months," and (3) "after 2 months."

3. You may find that you cannot try out your ideas exactly as you envisioned them or that it is difficult to be specific. That's o.k.--it is still important to write out your <u>intent</u>, as a tentative plan, knowing you may have to modify it once you are back on the job. <u>Try to develop at least two or three action items</u>; one may not work, so it's handy to have others available.

---

* Chart modified from <u>Interservice Procedures for Instructional Systems Development: Phase II--Design</u>, Robert Branson et al., Florida State University. AD-A019, p. 12.

- A target level of performance, projected without reference to the prior performance of the training population
- Pre-training performance of the training population
- Post-training performance of a nontraining population

Although the ideal measurement strategy may call for a combination of the latter two comparative bases--a pre-/post-training design with a trained (experimental) and suitable non-trained (control) group--this level of rigor is often unnecessary and/or infeasible. Thus, we specify comparative bases in the exhibits mainly for those few variables that are clearly not amenable to all three types of comparisons.

The information in the exhibits can be put to best use if these considerations are kept in mind:

- The exhibits do not replace general works on evaluation and training evaluation. Development of a sound evaluation design will require consultation of general works such as Campbell and Stanley, Goldstein, Kirkpatrick, and Peterson.
- Some evaluations may involve differentiation of the effects of input variables (e.g., the effects of prior trainee education or managerial experience) or program activities (e.g., participative training with and without experiential exercises) on knowledge and behavior. Because the realm of these differentiations is vast, we do not venture to specify them here.
- Because a useful evaluation design is one that is appropriate to the program evaluated, it is important to select evaluation measures carefully. This means that, from among the alternative measurements listed, measurements and comparisons should be selected that are necessary and sufficient to meet the levels of required rigor and confidence. This also means that, in selecting measures, it is important to determine how much agreement exists--among oversight and funding agencies, program directors, instructional and administrative personnel, trainees, chiefs, and training coordinators in sending agencies--about what it is that the program does and what it can reasonably expect to accomplish.
- It is essential to specify for a program the time-frames and range of acceptable values connected with desired results. For example, a program that merely proposes to "move departments toward more participative management" might be judged a failure for two apparently different reasons--because it had moved departments too little for one observer's tastes and too much for another's. In both cases, the program is judged a failure because it failed to specify the range within which change was considered acceptable.
- The variables and measures in the exhibits are meant to be only suggestive. The do not exhaustively state the variables in the models or the potential measurement options.

SELECTED BIBLIOGRAPHY

A. OTHER NATIONAL EVALUATION PROGRAM REPORTS ON POLICE MANAGEMENT TRAINING

Ross, J. G., Shanley, M. T., and Williams, J. L., Preliminary Conceptual Framework: National Evaluation Program Phase I Assessment of Police Management Training Programs. Washington, D.C.: Law Enforcement Assistance Administration, 1978. [Includes extensive bibliography; may be borrowed through interlibrary loan from the National Criminal Justice Reference Service.]

Ross, J. G., Shanley, M. T., and Williams, J. L., Summary Report: National Evaluation Program Phase I Assessment of Police Management Training Programs. Washington, D.C.: U.S. Government Printing Office, 1981.

Shanley, M. T., and Ross, J. G., Evaluation Report: Police Executive Development Institute (POLEX). Washington, D. C.: Law Enforcement Assistance Administration, 1980. [May be borrowed through interlibrary loan from the National Criminal Justice Reference Service.]

Shanley, M. T., and Ross, J. G., Evaluation Report: Line Supervisor Course [SEFICJ]. Washington, D.C.: Law Enforcement Assistance Administration, 1980. [May be borrowed through interlibrary loan from the National Criminal Justice Reference Service.]

B. NATIONAL STANDARDS REPORTS

American Bar Association Project on Standards for Criminal Justice, The Urban Police Function. New York: American Bar Association, 1972.

National Advisory Commission on Criminal Justice Standards and Goals, Police. Washington, D.C.: U.S. Government Printing Office, 1973.

National Advisory Commission on Criminal Justice Standards and Goals, Police Chief Executive--Report of the Police Chief Executive Committee of the International Association of Chiefs of Police. Washington, D.C.: U.S. Government Printing Office, 1976.

President's Commission on Law Enforcement and the Administration of Justice, Task Force Report: The Police. Washington, D.C.: U.S. Government Printing Office, 1967.

QUESTIONS ABOUT YOUR ACTION ITEMS

A. PRELIMINARY NATURE OF PLAN

1. Were you specific in writing the action item: (What will you need to do when you return to work in order to find out which actions are possible?)

B. RESOURCES

2. Who would be carrying out the proposed action, or helping with it (formally or informally)? Are the skills for carrying it out available? About how much time would this take?

3. Are there special materials or equipment required? What is involved in obtaining them? Will you be using a tool or system or aid from this course? How much adaptation is required?

4. Is continual monitoring or follow-through required? Who will do it?

C. IMPLEMENTATION

5. Do you have the authority to implement the action? If not, who does? How do you think you can go about getting approval?

6. What do you think the dgree of support is for your idea? Will you need to sell people on it? Who?

D. EFFECTS

7. Whom will this action affect? How will it affect them? Will anyone be the worse for the results? Anyone improved? What will be affected?

E. ENVIRONMENT

8. What in the organizational environment might interfere with your doing this? What in the organization would support your efforts?

C. MANAGEMENT AND POLICE MANAGEMENT

Baddeley, F., "National Police College of England and Wales," Journal of Criminal Law, Criminology, and Police Science, 63 (September 1972): 434-438.

Drucker, P. F., Management: Tasks, Responsibilities, Practices. New York: Harper & Row, 1974.

Ellis, L. A., "Administrator/Police Chief Relationship--View From the Administrator," Public Management, 56 (July 1974): 8-9.

Grosman, B., Police Command--Decisions and Discretion. Toronto: Macmillan, 1975.

Haimann, T., and Scott, W. G., Management in the Modern Organization. Boston: Houghton-Mifflin, 1974.

International City Management Association, Local Government Police Management. Washington, D.C., 1977.

Johnson, B. W., "The Manager and the Community's Chief of Police," Public Management, 56 (July 1974): 14-15.

Lee, R., "Administrators and Chiefs: Through a Looking Glass," Public Management, 56 (July 1974): 2-5.

Leonard, V. A., and Moore, H. W., Police Organization and Management. Mineola, New York: Foundation Press, 1971.

Likert, R., The Human Organization. New York: McGraw-Hill, 1967.

Likert, R., New Patterns in Management. New York: McGraw-Hill, 1961.

McClellan, J. E., "The Changing Nature of Police Management," Law and Order, May 1978: 18-25.

McGregor, D., The Human Side of Enterprise. New York: McGraw-Hill, 1960.

Marquis, D. R., "Selecting a Police Chief," Public Management, 56 (July 1974): 16-17.

Mayo, L. A., Analysis of the Role of the Police Chief. Unpublished doctoral dissertation, American University, 1980.

Reams, R., Kuykendall, J., and Burns D., "Police Management Systems--What Is an Appropriate Model?" Journal of Police Science and Administration, 3 (December 1975): 475-481.

Reichert, W. E., "Administrator/Police Chief Relationship--View From the Police Chief," Public Management, 56 (July 1974): 10-11.

Whisenand, P. M., and Ferguson, F. F., The Managing of Police Organizations. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Wilson, J. Q., Varieties of Police Behavior--The Management of Law and Order in Eight Communities. New York: Atheneum, 1970.

Wilson, O. W., and McLaren, R., Police Administration. New York: McGraw-Hill, 1972.

D. TRAINING

Auten, J. H., Training in the Small Department. Springfield, Illinois: Charles C Thomas, 1973.

California Commission on Peace Officer Standards and Training, A Final Report on Revision of the Executive Development Course. Sacramento, 1974.

California Commission on Peace Officer Standards and Training, A Final Report on Revision of the POST Supervisory and Management Courses. Sacramento, 1976.

"Change Through Training," FBI Law Enforcement Bulletin, 39 (July 1970): 2-4.

Ferry, J. D., and Kravitz, M. (eds.), Police Training: A Selected Bibliography. Washington, D.C.: U.S. Government Printing Office, 1980.

Grant, A., "Towards a Model for Police Management Training," Criminal Law Quarterly, 19 (September 1977): 291-302.

Hannon, M. J., "Building Blocks, Tinker Toys, and Role-Playing Games for Training Police Sergeants and Lieutenants," Law and Order, 20 (February 1972): 18-24.

Hartenstein, A. A., "A Cooperative Approach to Police Management Training," Journal of Police Science and Administration, 1 (December 1973): 433-439.

National Planning Association, National Manpower Survey--Volume 2: Law Enforcement. Washington, D.C.: U.S. Government Printing Office, 1978.

National Planning Association, National Manpower Survey--Volume 5: Education and Training. Washington, D.C.: U.S. Government Printing Office, 1978.

Neiderhoffer, A., Behind the Shield: The Police in Urban Society. Garden City, New York: Doubleday, 1967.

O'Neill, M. E., "Training--A Conceptual Model for Organizational Development," Police Chief, 40: 67-68.

Reinke, R. W., "Police Professionalism: A National Certification Program," Public Management, 56 (July 1974): 18-19.

Saunders, C. B., Upgrading the American Police. Washington, D.C.: Brookings Institution, 1970.

Shagory, G. E., and Deaneault, H. N., "Police Training in Managerial Budgeting: An Application of Andragogy," Journal of Police Science and Administration, 3 (September 1975): 327-331.

Strecher, V. (ed.), "Police Personnel and Training." Volume 13. In Cyclopedia of Policing: A Feasibility Study. Los Angeles: Loyola University of Los Angeles, 1973.

E. PROGRAM DEVELOPMENT AND EVALUATION

American Society for Training and Development, Training and Development Handbook: A Guide to Human Resource Development. New York: McGraw-Hill, 1976.

Anderson, B., Ball, S., and Murphy, T., Encyclopedia of Educational Evaluation: Concepts and Techniques For Evaluating Education and Training Programs. Washington, D.C.: Jossey-Bass, 1975.

Bollinger, S. and Vezner, K. O., Police Training Evaluation: A Systemic Approach. Toledo, Ohio: Toledo/Lucas County Criminal Justice Supervisory Council, 1975.

Campbell, D. T., and Stanley, J. C., Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

Goldstein, I. I., Training: Program Development and Evaluation. Monterey, California: Brooks/Cole Publishing Company, 1976.

Kirkpatrick, D. L. (ed.), Evaluating Training Programs. Madison, Wisconsin: American Society for Training and Development, 1975.

Peterson, R. O. (ed.) Determining the Payoff of Management Training. Madison, Wisconsin: American Society for Training and Development, 1979.

Ross, N. L., "Business School Ratings Studied." Washington Post (December 28, 1978): Section H, 1-2.

U.S. Office of Personnel Management. Report of the Training Evaluation Demonstration Project: Issues and Recommendations. Washington, D.C.: U.S. Government Printing Office, 1979.

Van Maanen, J., The Process of Program Evaluation. Washington, D.C.: National Training and Development Service, 1973.

F. PROGRAM EVALUATIONS

Cunningham, W. C., and Taylor, T. H., "Police Personnel and Training Evaluation," In How Well Does It Work? Review of Criminal Justice Evaluation 1978. Washington, D.C.: U.S. Government Printing Office, 1979.

Dunlap and Associates, Inc., Evaluation of CMPTC Recruit and Command/Supervisory Training. Hartford: Connecticut Justice Commission 1978.

Egger, Steven A., Third Year Evaluation of National Sheriffs' Institute. Washington, D.C.: National Sheriffs' Association, 1978.

Hettinger, B., Summative Evaluation of Cycle I, Regional Training Workshops. Washington, D.C.: University Research Corporation, 1977.

Hornaday, J., and Kaiser, R., The Effectiveness of a Managerial Training Program for Command Law Enforcement Officers. Wellesley, Massachusetts: New England Institute of Law Enforcement Management, 1972.

Newman, C. L., and Price, B. R., "Police Executive Development: An Educational Program at the Pennsylvania State University," Police Chief, 14 (April 1974): 74-77.

Newman, C. L., Price, B. R., and Horner, J. R., "Police Executive In-Service Training and Its Effect on Selected Personality Traits (Part 1)," Police Law Quarterly, Spring 1973: 14-27.

Newman, C. L., Price, B. R., and Horner, J. R., "Police Executive In-Service Training and Its Effect on Selected Personality Traits (Part 2)," Police Law Quarterly, Summer 1973: 42-47.

Price, B. R., and Adelberg, S., "Evaluation of Police Supervisory Training Using a Multivariate Assessment of Attitude Change," Journal of Police Science and Administration, 5 (March 1977): 69-73.

Rush, G. E., An Evaluation of Police Middle Management Training in California. Unpublished doctoral dissertation, Claremont Graduate School, 1975.

END