Implementation and Meta-Evaluation of an Experimental Method for Evaluating an Administrator Training Program

Nancy B. Larsen
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Educational Leadership Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/3253

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
IMPLEMENTATION AND META-EVALUATION OF AN EXPERIMENTAL METHOD FOR EVALUATING AN ADMINISTRATOR TRAINING PROGRAM

by

Nancy B. Larsen

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Education
Department of Educational Leadership

Western Michigan University
Kalamazoo, Michigan
December 1985
IMPLEMENTATION AND META-EVALUATION OF AN
EXPERIMENTAL METHOD FOR EVALUATING AN
ADMINISTRATOR TRAINING PROGRAM

Nancy B. Larsen, Ed.D.
Western Michigan University, 1985

The purpose of this study was to operationalize, test, and
assess the utility, feasibility, propriety, and accuracy of the
"success case method." A review of the literature revealed that
although evaluation is presently being conducted in educational
training programs, it is often not practical nor cost-effective. The
Kirkpatrick model for evaluation provided the conceptual framework
for viewing evaluation of training. A training managers' interview
questionnaire was developed to elicit perceived changes in trainees'
performance after training. Nine success cases, selected from a
group of 37 administrators, were interviewed by telephone 2 weeks
after the training. Selection criteria and procedures were developed
by the researcher. Interview information was analyzed and evaluation
reports were written and presented to several evaluation audiences.

Meta-evaluation of the success case method took place after
completion of the evaluation report writing. The Joint Committee
Standards (1981) provided a set of standards to assess the utility,
feasibility, propriety, and accuracy of the success case method as an
evaluation method. Judgments were made by the researcher based on
interviews with company training managers and three success case
trainees.
The major conclusions drawn from the study were:

1. The success case method was successfully operationalized and tested in the administrator training program.

2. The success case method was judged to be feasible for use by persons with little formal evaluation training. The method was cost-effective, and an accurate source of information for program revisions.

3. Limitations include the potential bias of self-reporting, the multiple roles played by training managers and the researcher, and interpretation of content analyses for report summaries.

The success case method was discussed in relation to practice and research in educational evaluation. Recommendations were made for potential users of the success case method.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. THE PROBLEM AND ITS BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Rationale for This Study</td>
<td>4</td>
</tr>
<tr>
<td>A Framework for Viewing Training Evaluation</td>
<td>6</td>
</tr>
<tr>
<td>The Purpose of the Study</td>
<td>8</td>
</tr>
<tr>
<td>Assumptions</td>
<td>10</td>
</tr>
<tr>
<td>Research Questions</td>
<td>10</td>
</tr>
<tr>
<td>Study Design</td>
<td>11</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>13</td>
</tr>
<tr>
<td>Summary of Current Issues and Problems</td>
<td>15</td>
</tr>
<tr>
<td>Overview of the Dissertation</td>
<td>15</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>17</td>
</tr>
<tr>
<td>A History of Evaluation</td>
<td>18</td>
</tr>
<tr>
<td>What Is Training Evaluation?</td>
<td>21</td>
</tr>
<tr>
<td>The Current Status of Training Evaluation</td>
<td>27</td>
</tr>
<tr>
<td>Methods of Inquiry</td>
<td>30</td>
</tr>
<tr>
<td>The Success Case Method</td>
<td>31</td>
</tr>
<tr>
<td>Research Designs</td>
<td>33</td>
</tr>
<tr>
<td>Ways of Using Standards for Evaluation</td>
<td>34</td>
</tr>
<tr>
<td>The Importance of Meta-evaluation</td>
<td>35</td>
</tr>
</tbody>
</table>
Table of Contents--Continued

CHAPTER

The Joint Committee Standards .................................. 36
Summary of the Chapter ........................................... 37

III. DESIGN AND METHODOLOGY ................................. 39

The Training Evaluation ........................................... 40
The Context of the Study ........................................ 40
The Administrator Training Program ............................ 41
Evaluation Questions ............................................. 43
The Selection Criteria ........................................... 43
The Selection of Success Case Trainees ......................... 44

Demographic Characteristics of the Population ............... 44
The Instrumentation .............................................. 45
The Administration of the Instrument ........................... 46
The Report of Findings From the SCM ........................... 47
The Information Analysis ........................................ 48

Meta-evaluation .................................................... 51
The Problems Encountered in Implementing the SCM .......... 51
The Strengths and Limitations of the SCM ..................... 51
Background of the Meta-evaluation ............................... 52
The Meta-evaluation Questions .................................. 52
The Rationale for Using Selected Standards for Making Judgments About the SCM .......................... 53

Meta-evaluation Data Collection ................................ 55
Analysis of Meta-evaluation Data ............................... 55
Table of Contents--Continued

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendations to Users of the SCM</td>
<td>56</td>
</tr>
<tr>
<td>Summary of the Chapter</td>
<td>56</td>
</tr>
<tr>
<td>IV. RESULTS OF THE STUDY</td>
<td>57</td>
</tr>
<tr>
<td>Sources of Meta-information</td>
<td>57</td>
</tr>
<tr>
<td>Interpretation</td>
<td>62</td>
</tr>
<tr>
<td>Question 1: Was the Information Provided by the SCM Useful?</td>
<td>63</td>
</tr>
<tr>
<td>Question 2: Was it Feasible to Implement the Success Case Method?</td>
<td>65</td>
</tr>
<tr>
<td>Question 3: Was the SCM Implemented in an Ethical Manner?</td>
<td>68</td>
</tr>
<tr>
<td>Question 4: Did the SCM Provide Accurate Information for Evaluation Uses?</td>
<td>69</td>
</tr>
<tr>
<td>Analysis of Meta-evaluation Data</td>
<td>73</td>
</tr>
<tr>
<td>Summary of the Chapter</td>
<td>73</td>
</tr>
<tr>
<td>V. CONCLUSIONS AND RECOMMENDATIONS</td>
<td>74</td>
</tr>
<tr>
<td>Summary of Evaluation Problems</td>
<td>74</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>75</td>
</tr>
<tr>
<td>Strengths of the SCM</td>
<td>77</td>
</tr>
<tr>
<td>Weaknesses of SCM in This Context</td>
<td>78</td>
</tr>
<tr>
<td>Recommendations to Users of SCM</td>
<td>79</td>
</tr>
<tr>
<td>Suggestions for Future Research</td>
<td>82</td>
</tr>
<tr>
<td>Summary</td>
<td>83</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>84</td>
</tr>
<tr>
<td>A. Summary of Joint Committee Standards</td>
<td>85</td>
</tr>
</tbody>
</table>
Table of Contents--Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. ATP Seminar Agenda</td>
<td>90</td>
</tr>
<tr>
<td>C. Evaluation Process</td>
<td>95</td>
</tr>
<tr>
<td>D. Demographics of Administrators</td>
<td>97</td>
</tr>
<tr>
<td>E. Demographics of Success Cases</td>
<td>99</td>
</tr>
<tr>
<td>F. Training Manager's Interview Form</td>
<td>101</td>
</tr>
<tr>
<td>G. Evaluation Highlights</td>
<td>107</td>
</tr>
<tr>
<td>H. Group Interview of Training Managers</td>
<td>113</td>
</tr>
</tbody>
</table>

BIBLIOGRAPHY                                                   | 116     |
LIST OF TABLES

1. Frequency Distributions .................................. 49
2. The Strengths and Limitations of the SCM .............. 71
LIST OF FIGURES

1. Evaluation Levels ........................................ 7
2. Timing of Meta-evaluation of Success Case Method .... 9
3. Sample Profile ............................................. 60
CHAPTER I

THE PROBLEM AND ITS BACKGROUND

Evaluation of training programs for management personnel is a topic of increasing concern to business and industry. The resources expended on administrator training is an indication of the perceived importance of training programs to top management. "Management expects training programs not only to have positive effects on the organization's accomplishments, but also not to cost more than those improvements are worth" (Salinger & Deming, 1982, p. 26). Many human resource developers either do not evaluate training programs or the evaluation process used is limited to participant reactions to the instructor or methodology. Worthen and Sanders (1973) stated, "Evaluation is one of the most widely discussed but little used processes in today's educational systems" (p. 1).

In order to determine the results of training and the impact of training on job performance, evaluation must go beyond the level of the immediate reactions of participants regarding the instructor or methodology. While it may be important to assess reactions, reactions do not in and of themselves "prove that the program will be translated into improved performance on the job" (Heyel, 1963, p. 961). Brandenberg (1982) observed that although numerous studies have focused on evaluation of training programs, there is no universally accepted model for evaluating training. There is, however, one generally accepted model which has been conceptualized by
Kirkpatrick (1971) to describe the evaluation process as it takes place at various times and stages of training. The Kirkpatrick model views evaluation as something more than just participants' reactions and changes resulting from learning. This model observes behavior changes on the job as well as long term impacts on the organization as a whole.

New approaches that are practical and useful in responding to the concerns and needs of management need to be tested. Warren (1979) stated, "The responsibility of the training function goes beyond the development of the training program; it must achieve the desired behavior changes in individuals or groups . . . and it must be measurable in terms of the organization's requirements" (p. 8).

Are there effective methods for measuring results of training? Can the methods be operationalized and tested? Is it possible to assess the utility, feasibility, propriety, and accuracy of evaluation methods? What can be said about the strengths and shortcomings of new methods? This study was stimulated by major questions such as these.

Statement of the Problem

Many of the current practices for evaluating human resource management training programs are costly and impractical. A common concern among company executives is that evaluation efforts should increase the potential for transfer of learning to the job. It should make programs more effective (Alkin, Daillak, & White, 1979). Training evaluators are increasingly being held accountable for the
quality and usefulness of their work. Participants' change in job behavior resulting from training is an important part of determining the impact that training has on the organization (Salinger & Deming, 1982). How to gather impact information that is useful is a problem which may exist, according to company executives.

Some efforts have been made to develop new evaluation methods to address previous shortcomings. One such method called the "success case method" (Brinkerhoff, 1983) has potential for addressing the problems of utility and practicality. The success case method focuses on assessing the performance changes and results of training. The purpose is to collect descriptive data about the uses and benefits of the training for certain participants. "Success case studies help explain how successful trainees make use of the training" (Brinkerhoff, 1983, p. 59). Other methods of inquiry exist for gathering data about results and impact of training (e.g., observations, work samples, and sales records), but these are typically expensive and time consuming. The purpose of the success case method is to gather formative data and "have significant value at little cost" (Brinkerhoff, p. 58).

The problem identified in this study was: How can methods for evaluating management training programs be improved? More specifically, what are some of the strengths and weaknesses of the success case method for evaluating an administrator training program?
Rationale for This Study

The human resource and development literature is filled with studies about training and evaluation. Many of these studies are aimed at the methodologies used in evaluation (Brandenberg, 1982; Braun, 1979; Brethower & Rummler, 1979; Brinkerhoff, 1983; and others), while a few have addressed the issue of utility and worth. The fact that training evaluation is not more useful is not due to a lack of attention to the topic, but rather because the reasons for evaluating are not clear. Springer (1979) stated that the potential purposes for evaluating management training are as diverse as the purposes for training. In addition, the array of potential purposes is influenced by the interested parties. Those parties may be the funding agency, the program designer, the trainer, future participants, bosses, and others. The evaluator must keep in mind that there are always multiple purposes and multiple clients for both training and evaluation of training.

It is important to design studies that evaluate new methods that are responsive to the needs of trainers and company executives, according to Dopyera and Dopyera (1980). The success of training programs should be assessed several weeks after training to measure results and impact. Further, Dopyera and Dopyera stated:

If evaluation is going to be worth the trouble of going beyond reaction forms, it must be more convincing than its past record has generally demonstrated . . . we must clearly consider alternatives to current methods for documenting and learning about long-term as well as side effects of training. (p. 67)
The challenge given by Dopyera and Dopyera (1980), Brandenberg (1982), and others in the field of human resource development and training evaluation suggests that development of new approaches for measuring the impact of training are needed. New approaches should have the capacity of the following: documenting relevant training applications objectively on the job; being responsive to concerns and issues of company trainers and company executives; and providing practical, efficient, ethical, and accurate information. Because assessing changes in job behaviors is more difficult, costly, and time-consuming than measuring reactions, it appears that human resource developers should be searching for viable alternatives. What is needed are evaluation methods that will "demonstrate empirically that training can make a difference" (Brinkerhoff, 1983, p. 58).

One question researchers are asking is: Why are trainers not doing more evaluation? In a recent article entitled "Practical Strategies for Evaluating Training," Salinger & Deming (1982) discussed several practical ways of answering the critical evaluation questions about the degree to which learning is transferred to the job and the impact it has on the organization. Salinger and Deming suggested some practical problems that indicated why training evaluation is not taking place:

1. Human resource development evaluators are limited by constraints on time, resources, and access to personnel for follow-up.

2. Current methods are not useful and practical.

3. Human resource development practitioners lack the background and training to conduct effective evaluations.
4. The nature of management training programs is such that the trainer cannot observe the trainee in order to measure effects of training.

5. Management wants results in a few weeks when evidence of organizational changes are long-term.

6. The appropriateness of each evaluation strategy depends on the characteristics of the training program, resources, and commitment from top management.

If the solution to this complex and costly problem of evaluating training requires an extensive background in evaluation methodology, then most trainers will not be successful. Salinger and Deming (1982) concluded that effective evaluation can still be conducted under far from ideal conditions, i.e., the conditions evaluators normally face in their organizations. In order to more fully understand the myriad of relationships between the functions of evaluation and the diverse needs of the organization, a framework for evaluating training has been suggested. Perhaps evaluators should put more effort into devising ways of making evaluation more practical and less effort into complaining about "far from ideal" conditions.

A Framework for Viewing Training Evaluation

Much controversy exists among trainers and evaluators concerning what they are trying to evaluate and why they are evaluating. The reasons for evaluating are often varied and unclear. "Everyone knows that there are benefits to be derived from evaluating training, nevertheless, confusion still exists" (Dopyera & Dopyera, 1980,
A conceptual framework that demonstrates the potential relationship between the training function and the fulfillment of the needs of organizations would be beneficial to human resource developers and company executives. The confusion exists because the evaluation process has not been separated into discrete elements for better understanding. Examination of the evaluation process from different levels provides the reader with a clearer understanding of the process. Figure 1 illustrates a familiar process model which utilizes categories or levels found in the studies of Brethower and Rummler (1979), Hamblin (1974), and Kirkpatrick (1975).

**EVALUATION LEVELS**

LEVEL 1—REACTIONS of Trainees

LEVEL 2—LEARNING Changes

LEVEL 3—BEHAVIOR on the Job

LEVEL 4—RESULTS to the Organization

Figure 1. Evaluation Levels.

For the purpose of understanding this study, the following definitions are helpful:

**Reactions** are defined as the stated feelings or opinions of the participants toward the quality of instruction.

**Learning** is defined as those skills, knowledge, and attitudes that training is expected to change.
Behavior is defined as the extent to which trainees applied new training skills on the job.

Results is defined as the effect new training skills had on the organization.

When a training program is completed, the trainer should be accountable for assessing not only how the program was carried out, but how the training program ultimately affected the organization in terms of the job performance of the participants. This conceptual framework provides the human resource development practitioner with a comprehensive perspective through which one is able to understand the evaluation process.

The Purpose of the Study

The purpose of this study was to assess how useful and practical the success case method is for evaluating an administrator training program in a public sector business. During Phase 1 (March 1983), the success case method was operationalized and tested in a training program of a Fortune 500 company.

The success case method is a "low-cost, high-yield" follow-up evaluation method described by Brinkerhoff (1983). It provides a means of exploring the changes which take place on the job as a direct result of training. Certain trainees who appeared to have benefited more from the training experience than most trainees were selected for intensive follow-up interviews. The trainees were interviewed several weeks after they had completed their training. The success case interviews focused on describing instances in which
trainees had applied new training skills on the job, benefits attributed to use of the training, and the problems encountered when using the training. The timing of the meta-evaluation of the success case method took place as Phase 2 of this study (see Figure 2).

![Figure 2. Timing of Meta-evaluation of Success Case Method.](image)

During Phase 2 a series of interviews took place with persons directly involved in using the success case method. The success case method was assessed using the standards of the Joint Committee on Standards for Educational Evaluation (1981) to make judgments in regard to its (see page 37 for complete definitions): (a) utility, (b) feasibility, (c) propriety, and (d) accuracy. These standards served as bench marks against which recommendations were rendered concerning the strengths and limitations of the success case method as an approach for evaluating administrator training programs. This
secondary evaluation is referred to as meta-evaluation: evaluation of an evaluation method.

Assumptions

The understanding of and rationale for this study rest upon the following assumptions:

1. Evaluation of management training programs will continue to be a major concern to human resource developers and decision makers.

2. For most trainers in the field, present evaluation methods for measuring results and benefits of training are both costly and complex.

3. The Joint Committee (1981) standards are a sufficiently developed and accepted definition of evaluation practice to be used to assess the success case method as it was applied in this study.

4. This specific administrator training program provided an adequate context in which to operationalize and field test the success case method.

Research Questions

This study was designed to address the following questions:

1. How was the success case method operationalized and tested within the context of a specific administrator training program?

2. What problems were encountered in implementing the method?

3. To what extent was the success case method effective and practical as an evaluation method?
4. What recommendations can be made for future users and uses of the success case method?

Study Design

During Phase 1 of this study the success case method for evaluating an administrator training program was operationalized. It was first necessary to define criteria and guidelines for the implementation of the success case method. This step included testing the method in a field setting. The setting for the test was a Fortune 500 company. The company had internally developed and implemented a new training program for recently promoted health care administrators. The success case method was operationalized to assess changes in job performance and additional benefits that resulted from attending a 1-week seminar in supervisory skills training.

The basic steps involved in the implementation of this method were:

Step 1: The researcher developed selection guidelines for "success cases," developed assessment instruments for evaluating the training, and designed data collection procedures for company trainers who conducted the success case interviews.

Step 2: The researcher established time lines and procedures for data collection to take place 4 weeks after training.

Step 3: The company trainers were trained to implement the success case method.

Step 4: The descriptive success case information gathered from trainees was recorded using a training manager's interview
instrument. Data were analyzed using a content analysis. Summary reports were written by the researcher.

Step 5: Presentations were made to various interested audiences, i.e., trainers, trainees, company managers, and executive officers.

The success case method interview instrument was designed by the researcher and reviewed by training managers to correlate the interview questions with the topics covered in the training program. Redundancies and omissions were corrected. The interview questionnaire was revised and used by company training managers several weeks after completion of the training experience. Findings were presented in written and oral presentations to audiences identified. These step-by-step procedures are amplified in Chapter III.

During Phase 2 of this research study a meta-evaluation was conducted to assess the success case method. The Joint Committee (1981) standards were applied to judge the utility, feasibility, propriety, and accuracy of the success case method as a method for evaluating training results. Several questions and subquestions were proposed by the researcher for each of the four categories of standards. Judgments were made by the researcher regarding strengths and limitations of this evaluation method in light of the standards used.

In order to confirm success case data and thus assess the accuracy standards, three persons were reinterviewed by the researcher after the information was analyzed and presented to the necessary evaluation audiences. These interviews took place with success case trainees who were interviewed earlier by their company training
managers, with training managers, and with company executives. The purpose of interviewing these different audiences was to assess the success case findings for accuracy, feasibility, propriety, and utility.

Three trainees from the program who were selected as success cases were reinterviewed by the researcher and asked if information gathered by company training managers was accurate, truthful, and representative of their perceptions of changes they said they had made on the job as a result of training. A company training manager was interviewed 10 months after training to determine how evaluation findings from success cases had been used to make decisions regarding further training efforts. The step-by-step procedures necessary to carry out this phase of the research design follow in Chapter III. Finally, conclusions were drawn and recommendations were made for potential uses of the success case method.

Significance of the Study

There is a need to field test and report on methods for evaluating on-the-job application of training programs. The evaluation researcher in particular has a responsibility for reporting the results found from testing innovative methods in terms of follow-up information which measures behavior changes and possible benefits and shortcomings of the training experience. Kirkpatrick (1975) supported this when he stated that reporting of results is: "A void which I hope my training and development colleagues will increasingly fill with training designs which produce measurable results--and that
they will then disseminate their evaluations through future articles" (Preface). Kirkpatrick stated that evaluation results could not be borrowed from other organizations but evaluation approaches and methods could (and should) be borrowed.

It is the hope of Kirkpatrick (1975), Brinkerhoff (1983), Dopyera and Dopyera (1980), and other evaluators that researchers test and evaluate practical approaches and methods which would fill the needs of companies, their training staff, and their top management. Evaluators and trainers should be stimulated to test methods which produce information of sufficient value to justify the resources expended. This study is important for the following reasons:

1. It can supply guidelines and caveats which others in the field of training evaluation might employ in using the success case method.

2. This study can stimulate further research and experimentation using the success case method and other methods in order to improve the quality and utility of training evaluation methods.

3. This description may serve to reduce some of the uncertainties involved in moving into unfamiliar territory in that it is presented by an evaluation practitioner with implications for other human resource development professionals.

While this innovative method is not all encompassing, i.e., success case method cannot answer all the questions there are about evaluation of training, it was operationalized; field tested; and evaluated for utility, feasibility, propriety, and accuracy in a training setting. Guidelines and applications for using the Joint
Committee (1981) standards to evaluate the strengths and limitations of the success case method may have impact on the field of training evaluation.

Summary of Current Issues and Problems

The problems in training evaluation identified in the literature can be summarized as follows: (a) often the evaluation methods currently in use are limited to reactions of trainees, (b) few evaluation methods look at the results and impact of training, (c) evaluation methods are not useful and feasible, (d) most methods for measuring trainees' job performance are costly and complex, and (e) few human resource trainers have the evaluation background necessary to carry out complex evaluation models.

There were no reported studies of meta-evaluation of an evaluation method. The current literature appears deficient in the area of meta-evaluation of innovative ideas and methods. Hopefully new research studies using standards for conducting meta-evaluation will be forthcoming.

Overview of the Dissertation

Chapter I is a statement of the problem, a rationale for the study, purpose of the study, research questions, study design, and significance of the study. Chapter II includes a review of the literature relative to evaluation training, methods, practice, issues, and problems. The chapter includes a conceptual framework for evaluating a comprehensive training program and the Joint
Committee (1981) standards for meta-evaluation of the success case method. Chapter III contains a description of the study design and methodology used, how the success case method was operationalized, the findings gathered using the success case method, standards which were useful in conducting a meta-evaluation of the success case method, and a summary. Chapter IV summarizes the findings and problems encountered which includes interpretations. Chapter V contains conclusions and recommendations.
CHAPTER II

REVIEW OF RELATED LITERATURE

The review of the literature focused on management level training and evaluation in books, research studies, and professional journals. The training and development journals provided a viewpoint relevant to practicing human resource trainers and evaluators, the context in which this research study took place. Many practitioners/trainers have experienced first hand the concerns associated with various evaluation methods and are likely to have insights into the practical problems of evaluation methodology.

Included in this review are nearly 60 studies found in computer searches conducted at the Western Michigan University Education Library and the Upjohn Corporate Business Library. Descriptors such as training, evaluation, management, evaluative research, and methodology were used to search the data bases of Lockheed and Bibliographic Retrieval Services (BRS) for abstracts and citations in Educational Resources Information Center (ERIC), Dissertation Abstracts International (Health Administration and Business Management Sections) from 1966 to 1984. The following questions guided the literature search: What is training evaluation? What problems exist? What has been tried and what was successful? What has not been successful? Are there needs for new evaluation methods? How can new evaluation methods be assessed?
A History of Evaluation

As early as the 1900s, Robert Thorndike was instrumental in providing the impetus for adopting measurement techniques to assess changes in individual learner behavior. The accreditation movement in education in the early 1900s also had an impact on evaluation methodology. Evaluation and measurement methods were focused on individual differences, as far as education was concerned, and had little relationship to improving school programs and curricula. Tests and measurements told something about individuals but nothing about the teaching of programs and curriculum. The schools had no reason to believe that the curriculum was not exactly what it should be. The accountability movement came later with the beginning of changes in the American culture—namely, scientific management. Top school officials were referred to as the "superintendents," the school building as the "plant," and the students as the "raw materials" (Guba & Lincoln, 1981). This trend was linked to the industrial revolution, success in science, and other aspects of the culture.

The first break with the earlier views came in 1932 with the writing of Ralph W. Tyler, a faculty member at The Ohio State University, who was named research director for the Eight Year Study. This study compared academic achievement of the students of 30 high schools across the U.S. (Madaus, Scriven, & Stufflebeam, 1983; Wentling, 1980). Tyler's main contribution to the field of evaluation was to insist that curricula needed to be organized around
certain educational objectives. Tyler made it clear that measurement and evaluation were separate processes with measurement being just a part of evaluation. His model of an evaluation process represented a major step in refining curriculum and educational programs although the model had certain limitations and weaknesses. It did not lead to judgment of merit or worth and did not provide for standards (Guba & Lincoln, 1981). The Tyler model provided educators with the concept of educational objectives to determine essential changes in human beings, but he was not clear about making judgments of value. The model did not provide guidance on how to evaluate data or how to invoke values—whose values or by what standards should performance be judged. By the late 1960s, Tyler had initiated the National Education Assessment Project (NEAP) for which he sampled thousands of students to compare academic performance.

With the Russian launch of Sputnik in 1957 came dramatic changes in the American educational funding resources. Until 1963, most evaluation studies were the result of federal or private funding projects and did not reflect the evaluative efforts of other educational agencies or individual states. A serious challenge to Tyler's model for evaluating objectives came when other evaluators made the point that evaluation needed to focus on decisions and decision makers rather than on achievement of objectives.

In 1963, Congress passed the Vocational Education Act which required that each state be responsible for evaluation of educational programs and projects. It was important to the government and Congress that the impact of new educational curriculum activities be
assessed. Unfortunately, most states did not immediately assume the responsibility for evaluation. In 1965, the Congress enacted the Elementary and Secondary Education Act (ESEA). This act required that all projects conducted under Titles I and III have an evaluation component.

As a result of mandated evaluation, new theories and papers criticizing the inadequacies of the early evaluation theorists appeared. New models emerged in response to concerns and issues raised. Early papers came from educational evaluators and included such topics as: goal-free evaluation (Eisner, 1969; Scriven, 1972, 1974); responsive evaluation (Stake, 1975); meta-evaluation (Stufflebeam, 1974); and naturalistic inquiry (Guba & Lincoln, 1981). These responsive approaches were an attempt to increase the utility and the credibility of evaluation findings.

The field of evaluation began to emerge as a distinct profession related to research, development, and evaluation of educational programs in the early 1970s. The major contributions in contemporary literature regarding program and project evaluation appeared in writings in journals, monographs, and networks, including Educational Evaluation and Policy Analysis, New Directions for Program Evaluation, Evaluation News, and others which deal exclusively with evaluation.

At the present time, many large universities have expanded their graduate curriculum to include courses in evaluation methodology. A select few universities—such as the University of Illinois, UCLA, and Western Michigan University—have certified graduate programs in
evaluation. Centers have been established at these universities which offer workshops and institutes on various evaluation topics. The state of Louisiana has a program for certifying evaluators.

There appears to be a need for expanding efforts to educate evaluators and practicing trainers about new evaluation methods. Although the subject of training evaluation in business and industry is often a topic of concern, "Increasingly, the field has looked to meta evaluation (Scriven, 1975; Stufflebeam, 1978) as a means of assuring and checking the quality of evaluations" (Madaus et al., 1983, p. 16). A joint committee (The Joint Committee on Standards for Educational Evaluation, 1981) consisting of 12 persons from professional organizations has issued a comprehensive set of standards for judging evaluation and educational programs. With the exception of Madaus et al. (1983), Brinkerhoff (1983), and Stufflebeam (1974), few writers provide guidelines and suggestions for conducting meta-evaluation using the Joint Committee standards. Although the gains of the past 15 years have been impressive, there are still many deficiencies. Leaders in the evaluation profession are beginning to apply the principles of educational evaluation and the principles of training to improve the utility of evaluation findings and to be responsive to the needs of clients.

What Is Training Evaluation?

Early research in the theory of training evaluation is discussed by Kirkpatrick (1975), Hamblin (1974), Hyman and Wright (1966), Hayes and Williams (1975), and others. Two early trends appeared in
evaluation research according to these researchers. First, evaluation had come to be expected as a regular part of training and development programs in business and industry. And second, there was a movement toward demanding more rigorous and objective evidence of success. "If we are to evaluate training (other than on the basis of pure hunch) we must collect some information about the changes that have been caused by the training" (Hamblin, 1974, p. 14). In order to become more objective and systematic about evaluation, it is important to think about what effect training will have and the sequence in which these effects will occur. Several kinds of evaluation are needed to provide the data needed to measure the effectiveness of training efforts. By separating the evaluation process into small components, evaluation methods, questions, functions, timing, and criteria can be developed and tested to assess desired results. Provided in the research findings of Kirkpatrick (1975), Hamblin (1974), and Brethower and Rummler (1979) is a conceptual framework to systematically view the training evaluation process during and after the training activity. That framework is referred to as the "levels of evaluation," with each level becoming more complex and difficult to evaluate than the previous. The levels and evaluation questions of the conceptual framework appear in Figure 1, page 7.

These levels of the evaluation process expose a number of dimensions which might be investigated to determine whether training is having the desired effect. Each level asks a different set of questions, uses different methods, each with their separate data sources and criteria.
EVALUATION LEVELS

Level 1--Trainee's REACTION

Level 2--The change in LEARNING

Level 3--BEHAVIOR change on the job

Level 4--The RESULTS to the organization

Level 1: REACTIONS asks the question, do trainees like the training? This level reflects the feelings of the participants toward the quality of instruction. Measurement and data analysis take place during or at the end of the program using questionnaires with a Likert-type scale, postcourse interviews or expectations evaluation.

Level 2: LEARNING asks the question, did trainees learn the skills, knowledge, and attitudes that training is expected to change? This level is generally assessed using a written test given before and after training, using a standardized attitude/aptitude questionnaire (either open or closed-ended) or by task analysis appropriate to the job.

Level 3: BEHAVIOR change asks the question, did the trainees apply what they have learned to change their performance on the job? Measurement takes place several weeks after training in a field setting. Questionnaires, surveys, observation, critical incidents, self-report, and interviews are typically used to gather feedback from trainees, their subordinates, or their superiors.
Level 4: RESULTS asks the question, does the training impact the organization over the long term? This question is answered many months after training by analyzing work flow records, activity sampling, cost-benefit analysis, and indices pertaining to sales or service.

Hamblin (1974) and Kirkpatrick (1975) said that trainers generally conduct evaluation at only the first two levels. They identify the reactions of the trainees to the training program. Sometimes the amount of learning the training program attempted to accomplish is measured through tests of attitudes or skills.

We seldom find evaluation conducted at the third or the fourth levels, because it takes time, effort (expertise) and money to conduct evaluation at these levels, and trainers are either too busy or too shortsighted to see the value of this investment. (Hamblin, 1974, p. 137)

A comprehensive research study conducted by Catalanello and Kirkpatrick (1975) reviewed the status of training and evaluation in business and industry. This study selected 154 companies who were presently using human resource development programs. Questionnaires were sent to these companies located throughout the U.S. and Canada. The response return was 71% (110). Company training managers used the Kirkpatrick conceptual framework to determine the extent to which evaluation was currently taking place. The findings showed that 78% of the companies had attempted to measure trainees reactions—Level 1. Typical methods were used to measure learning, and 21 companies attempted to measure on-the-job behaviors, while only 12 companies measured behavior before training. The most frequently used approach to measurement of behavior change was the interview method. Timing
of the interviews was an important variable according to the authors. Most companies attempted to measure behavior change within 6 months after completion of training. Conclusions were: The development of evaluation methodology is in the early stage; most companies are measuring only reactions to training programs; few studies are being conducted to measure evaluation Levels 2, 3, and 4; and few companies are attempting to measure results and impact of training on the company. The researchers concluded that measurement at the higher levels of evaluation requires time, money, and resources not always available in the everyday world.

A similar study to assess supervisory training needs and to determine the effectiveness within a national forest system was conducted by Braun (1979). Eighteen trainees responded to a questionnaire and 13 supervisors of trainees responded. One topic of focus was the training effectiveness as it improved organizational production. The researchers concluded they were unable to state whether training improvements made any difference because only 20% of the trainees responded to a mailed questionnaire and they could only make an educated guess as to worth or benefits. Conclusions were:

Follow-up by questionnaire alone is hazardous to the success of an evaluation effort. A more successful approach would have been a series of personal interviews with individuals or small groups of trainees (to discuss specific accomplishments and problems). (p. 10)

Further, many nonprofit and government organizations have experienced difficulties in developing reliable and valid effectiveness and benefit measures which can also measure individual contributions to productivity. Measuring training effectiveness in managerial and supervisory areas therefore is difficult. (pp. 8-9)
Further studies by Kirkpatrick and Probst (cited in Kirkpatrick, 1975) at the University of Wisconsin attempted to measure change in behavior by interviewing trainees and their bosses after a 3-month delay following a 3-day "supervisory skills seminar." They used a structured interview form to relate behavior changes to specific course content. The results "indicated wide variations in behavior change" (p. 99). Some supervisors had changed considerably and others hadn't changed at all. Researchers concluded: "It is obvious that very few training planners and coordinators have the knowledge, ability and time to evaluate in terms of behavior changes" (p. 99).

In an analysis of evaluative research studies already done, Mann (1972) compared 181 evaluative studies by using the criteria, "findings that had a greater degree of social significance and a lesser degree of experimental error than other research in their area" (p. 269). His sample represented information about the success of evaluative research to gain understanding about the process and methods of changing behavior most efficiently. Several comparative analyses were used in this study on variables, such as, aspects of experimental design, nature of the samples, instruments used to measure change, number of types of instrument used per study, errors of methodology, and the relationship between demonstrated change and methodological deficiency. In many cases no relationship existed between variables, negative relationships occurred, or positive findings were few. Two conclusions were suggested by the analysis:

The quality of evaluative research is remarkably poor, and there is little difference in results of evaluative studies conducted in different content areas. Specifically, there
is no indication that the findings of evaluative research are influenced by the method tested, the content area in which the test is conducted, the change criteria used, or the methodology quality of the study which the evaluation is made. The only clear positive finding is that change is demonstrated in approximately 45% of the studies. (p. 278)

The Mann (1972) study typically represents a sound research design which produced information of limited value and utility. An experimental study as complex as this is costly and may offer limited information relevant to the improvement of training.

Weiss (1972) also addressed the issue of the political problems facing the evaluator today. During the early years of evaluation when federal funding was dependent upon evidence of positive changes, many evaluators had reason to question evaluation practices and methods. Elinson (1972) reviewed the results of 10 of the most competent and best known published evaluations and found limited successful evaluation results. In another illustration by Weiss (1972), the writer stated that Ward and Kassebaum (1972) reported the lack of conclusive results of group counseling in a correctional institution. The institution's response to the study was to expand the program and dismiss the researchers.

The Current Status of Training Evaluation

Although the subject of training evaluation in business and industry is a topic of professional conversation and often appears in the journals, rarely does one find agreement as to what are the best evaluation methods or activities, according to Brandenberg (1982). In his recent study, Brandenberg attempted to determine what
evaluation practices are currently in use. Human resource trainers were asked about functions, roles, techniques, and skills needed in the field of training evaluation. Information was gathered by a questionnaire mailed to 50 training professionals. The sample was randomly selected into two groups and the return rate was 64% or greater. Most frequently used measurement methods were those methods administered "during" training activities. The methods used were considered "as efficient, short and less labor-intensive" (p. 17). Complex methods requiring more time or personnel received lower priority and were less often used. The least frequently used methods were measures of job performance and follow-up after training. Conclusions of the study were: Educational evaluation is firmly established and practiced in postsecondary settings. Business and industry training evaluation tends to overplay the importance of cost benefits, trainers seem to lack the necessary skills for conducting evaluations, and data collection techniques are rarely longitudinal in nature.

Recent studies indicate there may be limited relationship between training efforts and job performance. Gregory (1983), Loewy (1983), and Sampson (1981) conducted studies to measure employee performance changes after completion of a specific training program. Gregory's (1983) study showed no significant changes occurred in the workplace. He used both quantitative data and qualitative descriptions to sample 100 employees rating their managers after training. The researcher concluded: "Data from the qualitative interviews enhanced the meaningfulness of the questionnaire data and provided
The Sampson (1981) study was designed to develop an evaluation method and test it by determining the relationship between the training program for correctional officers and performance on the job. Sampson used the Kirkpatrick (1975) model for evaluating different levels of training effects and the results showed no significant relationship existed between training and job performance. No further conclusions were reported in this study.

Brandenberg (1982), in his discussion of the current status of training evaluation, stated his primary conclusions were: "Skills and techniques used in training evaluation need to be expanded and updated" (p. 14). Managers of training need to have a better understanding of evaluation in order to provide better quality courses and programs. Brandenberg suggested making follow-up evaluation easier and more useful.

Evaluating behavior changes on the job is a higher level task in Kirkpatrick's (1975) conceptual framework. Ideally, the behavior to be evaluated can be isolated and coded through observation or interview. This means that the training activity can be identified as separate from other behaviors, according to Warren (1979). For instance, if an administrator is to be evaluated, the problem becomes complex since the total performance of that administrator is relative to past experiences as well as to training activities. Warren (1979) addressed this issue head on when he said:

Before we can attempt to change behavior the required behavior must be known, and the means to measure the change in behavior must be found. Until these two constraints are
satisfied, the effectiveness of training actions cannot be known and their functional ability becomes guesswork. Behavior change must be measured in the actual job environment. It is performance on the job that is important, not performance in the training program. (p. 8)

Warren concluded that the only valid result of training experience is a measurable increase or improvement in the individual's contribution to the organizational goals. How those contributions are measured was unclear in his writings.

Methods of Inquiry

Selection of an evaluation method requires careful consideration. "The selection of a method is ultimately measured against existing standards, as the success of an evaluation is measured mostly by the success of the changes that occur with . . . training programs as a result of that evaluation" (Wentling, 1980, p. 3). Methodology strategies are sometimes best when using a combination of methods to study a program or project. In the case of evaluation research, this can mean using both quantitative methods and qualitative methods to study the same program. Denzin (1978) explained the logic of methodological "triangulation" as the use of multiple methods to study a single phenomenon. Triangulation is based on the premise that "no single method ever adequately solves the problem of rival causal factors . . . because each method reveals different aspects of empirical reality, multiple methods of observations must be employed. This is termed triangulation" (p. 28).

One segment of the training evaluation literature is concerned with what is called the "scientific" approach to evaluation.
Unfortunately, in the case of evaluation of training it is not always practical or feasible to prove or disprove a hypothesis about how effective training activities operate. Often more open-ended approaches to evaluation are used simply to "find out about things." This is an alternative method to the scientific approach and depends on the concerns and needs of decision makers. An alternative method to evaluating training programs is described in the literature by Brinkerhoff (1983) as the "success case method." This method is appropriate when conventional experimental methods are limiting.

The Success Case Method

The choice of evaluation methods, designs, and analysis strategies depends upon the key evaluation questions, resources available, time available, background of the evaluator, and the needs of the decision makers. An alternative method for evaluating training programs is described by Brinkerhoff (1983) as the "success case method." The success case method (SCM) for evaluating training is described as a low-cost and high-yield method for collecting descriptive information about how trainees have changed since training and what benefits they have received. "Success cases may be studied briefly by short interviews with a few trainees or through more elaborate approaches, i.e., task planning sheets, work analysis, or cost effectiveness" (Brinkerhoff, 1983, p. 58).

There are several uses for the success case method in training and development efforts such as those "which do not lend themselves to traditional quantitative measurement based evaluations"
(Brinkerhoff, 1983, p. 58). Examples of uses might include: new training programs still in the development stages, management training programs that have hard-to-measure benefits, or programs that have limited time and resources to conduct broad-scale assessment. In such cases, an evaluation method such as success case has potential value and is practical.

Shortly after completion of a specific training program, the evaluator would select a few trainees from the training program who seemed to benefit most from the experience. The selection procedure is not rigorous or scientific by nature, but rather an intuitive approach. Judgments are made relative to the persons who learned the content of the training program best, who were positive and contributed to the discussions, who were most likely to apply newly learned skills and knowledges, and who believed in the value of the training experience. Follow-up interviews may take place several weeks after completion of the training. Several key questions would guide the interviewer:

1. How have trainees used the training?
2. What benefits can be attributed to the use of training?
3. What problems were encountered in using the training skills?
4. Were there negative consequences of the training?
5. What criteria were used to decide if training skills were being used correctly?

It should be noted that SCM is not intended to gather testimonials. Rather, its purpose is to gather specific detailed and documented instances of training applications. The purpose of the
follow-up interviews with a few select trainees is to assemble specific documented descriptive information about training applications which will assist decision makers. This method differs from the more traditional experimental design with random sampling of large numbers. Purposive selection of a few cases makes success case method cost-effective, rapid, and simple when compared to more elaborate and costly evaluation methods. Descriptive information organized about specific key questions can be retrieved and coded quickly for presentation to company officials. One limitation to the case study method of inquiry is the problem of coding behaviors by categories and, secondly, analyzing the data in an objective manner.

Research Designs

By contrast, the pure experimental approach to evaluation includes random selection of participants from a target population. An experimental and control group is generally identified and data collected and compared. Data are often collected by means of standardized tests and comparative analysis conducted to determine significant differences. An alternative approach is some form of qualitative measurements to observe or inquire about important behaviors and events in a given population. If the evaluator is looking for interactions that may occur fortuitously, in-depth interviews conducted after training can offer descriptions useful for making revisions and decisions.

Experimental and quasi-experimental designs for collecting data often make quantitative comparisons which can be statistically
Variables can more easily be isolated and hypotheses tested. Different from quantitative information about a study are the qualitative measures which often use a discovery approach to uncover emerging patterns and trends. A variety of measures can be employed including questionnaires, survey instruments, observational checklists, or telephone interviews. Many times the qualitative approaches become a mixture of methods based upon the choices made by the evaluator and key audiences for the evaluation reports. One strength of the SCM is the capacity to collect descriptive information about a specific training program by using the telephone survey method.

Behavior changes are often difficult to observe, record, and interpret unless some predetermined standards or criteria have been established.

**Ways of Using Standards for Evaluation**

Suggestions for employing the standards-setting process are addressed by Ridings (1980), the Joint Committee on Standards (1981), Madaus et al. (1983), and Brinkerhoff, Brethower, Hluchyj, and Nowakowski (1983). "Possibly the most controversial aspect of professional standard setting revolves around the way standards are used or misused" (Ridings, 1980, p. 225). The Joint Committee (1981) has indicated in the introduction to the Standards for Evaluations of Educational Programs, Projects, and Materials that standards can be used as guiding principles and as a basis for self-regulation and accountability. Additionally, they have provided a series of
situations for which the standards may be applicable; among the functions are "training and educating evaluators, guiding and educating those who commission and administer evaluation and providing criteria for those who evaluate evaluations" (Ridings, p. 225). The Joint Committee (1981) has not yet formulated definitive guidelines for use of the standards.

Madaus et al. (1983) developed a matrix which analyzed the relative importance of the 30 standards in performing evaluation tasks. The matrix is useful in determining which criteria might apply to relevant questions asked. The Joint Committee (1981) has also developed a checklist form which describes different levels of use and/or consideration of the standards. This form provides a range of options for indicating the extent to which each standard was considered.

The Importance of Meta-evaluation

During the past 10 years evaluators have attempted to demonstrate that they are doing quality work (Stufflebeam, 1974). "The literature on evaluation provides limited guidance for evaluating evaluation work" (p. 5). The term meta-evaluation was introduced by Scriven (1972). Lessinger (1970), Provus (1971), and others have discussed the concept of educational auditing. Stufflebeam suggested that APA technical standards for tests and Buros's Mental Measurement Yearbooks (1965) are useful meta-evaluation tools. Also, Campbell and Stanley (1966) have published a set of standards commonly used for evaluating experimental and quasi-experimental research designs.
How does one apply a set of meta-evaluation criteria? When the appropriate set of meta-evaluation criteria has been selected, it may be necessary to make adaptations which will fit a specific situation. The first task is to determine how to measure compliance with each standard selected. Decisions can range from a dichotomous level, i.e., met or not met, to a rank order on a scale of 1 to 10. To determine compliance indicators, "you should consider the amount of information available, the level of precision needed, and the level of agreement possible" (Brinkerhoff et al., 1983, p. 218). The authors suggested possible methods to establish criteria for meta-evaluation and gave examples of different measurement approaches.

Brinkerhoff et al. (1983) also provided suggestions for conducting the meta-evaluation. Some of the suggestions include hiring a consultant, using a review panel or advisory group, and reviewing final reports. Meta-evaluation should safeguard the rights of persons involved, should be realistic, should be technically sound, and should be clearly useful.

The Joint Committee Standards

The Joint Committee (1981) standards for educational evaluation were prepared by a select committee concerned with evaluation standards consistent with good practice. Appendix A is a summary of the 30 standards grouped according to the four major categories essential to good evaluation: (a) utility, (b) feasibility, (c) propriety, and (d) accuracy. These standards provide the guiding principles for conducting meta-evaluation of evaluation methods useful to training
programs.

The "utility" standards reflect the consensus found in the literature and among constituent groups regarding the need for evaluations to be responsive to the needs of clients. This category contains standards for assuring that evaluation information is informative, timely, and influential.

The second set of standards is concerned with "feasibility" and addresses the issue that evaluation procedures must be cost-effective and workable in practical ways. Overall, the feasibility standards require evaluation to be realistic, prudent, diplomatic, politically viable, and frugal.

The third set of standards addresses "propriety" and reflects the fact that evaluation should serve the interests of many people in different ways. This category of standards is meant to protect the rights of persons affected by the evaluation. The propriety standards state that the evaluation must be conducted legally, ethically, and with due regard for the welfare of others.

The "accuracy" standards include standards that assure accuracy of information produced. These standards require that the information obtained is technically sound and that conclusions are linked logically with the data. The accuracy standards are an index of the overall validity of the evaluation.

Summary of the Chapter

Chapter II is a review of literature relative to the history and background of educational evaluation and the impact on training
evaluation. A framework for viewing evaluation of training was taken from the Kirkpatrick (1975) model. The status of evaluation methods and approaches was discussed, the success case method was described, and the standards for conducting meta-evaluation were described.

There appears to be an expanding effort to improve the quality of training evaluation designs and evaluation methods. Despite the impressive gains made in the past 15 years, evaluation of training is generally in the infancy stages. Although the Kirkpatrick (1975) model is widely discussed in the literature, few practitioners in the training field are using this model to evaluate behavior changes and impact on the organization. Many research studies infer that elaborate follow-up research is too costly or often the findings are dismissed as irrelevant to decision makers. The literature suggests that new methods would be helpful if they were easy to implement and were low-cost. A problem that exists is that many evaluations are dismissed as too weak methodologically and many other methods are dismissed as too costly. Needed are methods that answer a limited set of useful questions at low cost in time and dollars. Thus, to achieve good evaluation economically the human resource evaluator must limit the scope.
CHAPTER III

DESIGN AND METHODOLOGY

The purpose of this study was to operationalize, test, and evaluate the success case method for evaluating an administrator training program. Judgments were made in regard to the potential uses of the SCM to provide useful and accurate information in a timely and ethical manner.

The design and methods used in this study are organized around the two major events of the study—the training evaluation (which used a success case approach) and the success case method meta-evaluation. Research Question 1 is concerned with the environment in which training occurred. Research Questions 2, 3, and 4 are concerned with the meta-evaluation of the success case method.

**Research Question 1:** How was the success case method operationalized and tested within the context of a specific administrator training program? Issues examined for this question included:

1. Context of the study.
2. Administrator Training Program (ATP).
3. Evaluation questions.
4. Selection criteria.
5. Selection of success case trainees.
6. Instrumentation.
7. Administration of the instrument.
9. Information analysis.

Research Question 2: What problems were encountered in implementing the method? Issues examined for this question included:

1. Problems encountered in implementing the SCM.
2. Strengths and limitations of the SCM.

Research Question 3: To what extent was the SCM effective and practical as an evaluation method? Issues examined for this question included:

1. Background of the meta-evaluation.
3. Rationale for using 10 selected standards for making judgments about the SCM.
6. Analyses of meta-evaluation data.

Research Question 4: What recommendations can be made for future users and uses of the SCM? Issues examined for this question included.

1. Recommendations to users of the SCM based on the strengths and limitations of this study.

The Training Evaluation

The Context of the Study

The success case method was tested in a management training program for a Fortune 500 company. At the time of the study, the
company employed approximately 2,000 staff persons and several thousand home health nurses, aides, and homemakers. Each of the 200 or more service offices was staffed by one health care administrator and support staff. In addition, the home office employed approximately 50 management level personnel with support staff. The company training staff consisted of one training director, one evaluation consultant, and three training managers (one for each of three divisions, east, west, and central). Because this was the first major thrust to train administrative level personnel, the evaluation of the program was formative (evaluation used to improve an object while it is still being developed) rather than summative (evaluation designed to present conclusions about merit or worth). The training staff was responsible for delivery and evaluation of all management level training programs within their respective divisions. The training managers had previous training in conducting objective interviews. This was helpful background for conducting success case interviews.

The Administrator Training Program

The success case method was implemented during the follow-up segment of a newly developed administrator training program (ATP) designed for field office administrators. Seminar 1 of the ATP began in March 1983 with 37 company administrators. The ATP consisted of four 1-week seminars offered 4 to 6 weeks apart. The success case method was one of several methods used to evaluate Seminar 1, "Supervisory Skills Training Seminar." This training program was designed by the home office training staff and was delivered by three company
training managers.

The purpose of Seminar 1 was to upgrade present job skills in preparation for the increased leadership responsibilities of the administrators. Each of the three divisions was allotted funding for 15 trainees. The first group of trainees was selected by the company regional managers. Regional managers were provided with standardized guidelines and criteria for selecting trainees. Trainees were selected on the basis of "the best qualified person with the highest percentage of earnings over budget for 1982." This was an arbitrary decision set by management over which the researcher had no control. Further, management understood that this was not an immediate expectation of training. Administrator trainees were flown to the home office in Michigan for each of the four 1-week seminars. Curriculum content and methods for the training seminars were developed by the training director in consultation with advisory groups from the field and company executive officers. The six content modules offered in Seminar 1 were: (a) role of the supervisor, (b) problem solving/decision making, (c) leadership, (d) goal setting, (e) training, and (f) communications. (See Appendix B for the seminar agenda.) Company training managers received 3 days of training prior to each seminar. They were responsible for delivery of training for their divisions, as well as for administering all evaluation instruments. The evaluation consultant was under contract to design the evaluation procedures, assessment instruments, collection of data, analysis of data, and report writing in order to carry out the evaluation plan.
Evaluation Questions

The evaluation of Seminar 1, "Supervisory Skills" training, included assessment before training, during training, and after training. (See Appendix C for the overall evaluation design.) The success case method was one of several evaluation methods used to gather information of select trainees. The overall company concerns to be addressed by the evaluation design were determined by the evaluation consultant and the training director based on the Kirkpatrick framework for evaluating training (see page 23, Chapter II). These concerns were:

Level 1--Reactions. "Do trainees 'like' the training? How did they value the instruction?" as measured by daily reaction sheets.

Level 2--Learning. "Did trainees learn the skills, knowledge, and attitudes that training was expected to change?" as measured by a true/false pre- and posttest.

Level 3--Behavior. "Did administrators apply what they learned to change their performance on the job?" as measured by success case interviews.

Level 4--Results. "Did the training have impact on the organization? Did it change sales, profits, morale?" as measured by interviews with company managers and executive officers.

The Selection Criteria

Training managers were asked to select, on the basis of their perceptions, three success case trainees during the last day of
training. The following criteria were provided to training managers to assist in selecting the three success cases. "Success cases" are persons who: (a) learned the content of the training better than most, (b) were more positive and contributed more than others during discussion periods, (c) were more likely to apply the skills and knowledge taught in training, and (d) believed in the utility and worth of their training experiences. Nine trainees were selected and became the respondents for the success case interviews conducted by the training managers.

The Selection of Success Case Trainees

The sample of nine success case trainees was selected from the population of 37 health care administrators who participated in Seminar 1. The success case sample included three trainees from the east division, three trainees from the west division, and three trainees from the central division. Each training manager selected and interviewed three success cases from their respective division.

Demographic Characteristics of the Population

Demographic information about the 37 administrator trainees (Appendices D and E) is helpful when describing the context in which the SCM was tested. The total population consisted of 36 females and one male. Fifteen of the trainees were between the ages of 25 and 45. Twelve trainees were over 46 years of age, most were married (26), had been with the company less than 8 years, and had been in their present position less than 5 years. All 37 trainees had high
school educations, 12 had college or graduate degrees, and 14 had their registered nurse's degrees.

Of the nine success case trainees, eight were females. Their ages ranged from 26 to 55; eight were married. One was new to the company; the other eight were with the company between 1 and 8 years. Eight of the nine success cases had high school diplomas, while four had college degrees. The nine had been in their present positions 1 to 8 years.

The 37 administrator trainees were drawn from a pool of about 200 company administrators identified by their superiors as "successful administrators." Success cases were selected from this pool of "successful administrators." None of the nine success cases had been with the company more than 8 years. The demographic characteristics of the population and the sample selected appear in Appendices D and E. The success cases interviewed were selected by their training managers on the basis of several criteria developed by the evaluation consultant (see page 43). The purpose of this sampling strategy was to efficiently gather information from the most successful trainees. As suggested by Brinkerhoff et al. (1983), this type of purposive sampling based on successful cases can be most cost effective and useful when time does not permit other broad-scale assessment efforts.

The Instrumentation

The interview questionnaire was developed by the researcher and included 22 open-ended questions based on the goals and objectives of
the six content modules in Seminar 1. Draft copies of the interview questionnaire were reviewed by two evaluation practitioners and their recommendations were used to improve the format and style. A second draft was presented to the training managers to eliminate redundancies and to verify that the content of the seminar was fully represented in the questionnaire. The final draft was given to the training managers for use in conducting success case interviews by telephone. The telephone interviews were self-reports that focused on activities, behaviors, and attitude changes of the trainees. The Training Manager's Interview Form is in Appendix F.

Although the questions were open-ended, they included sub-questions which guided the interviewers and provided for consistency across interviews when administering the questionnaire. The items were based on the content of the "Supervisory Skills" Seminar 1. A sample question from the "Role of the Supervisor" module was, "Did you encounter problems when you delegated a job? If so, what should have been done?"

The Administration of the Instrument

Telephone interviews with the nine success cases were conducted by the training managers 2 to 3 weeks after training. Training managers telephoned the nine trainees to schedule interview appointments and explain the purpose for the interview. Each person selected agreed to volunteer. Before interviews, the training managers were given detailed instructions to avoid leading questions, to ask about actual experiences, to ask for examples of behavior change,
and to ask the "why" behind the actions. The decision was made by the training director to have the training managers administer all evaluation instruments. This provided the trainers with ongoing commitment to the evaluation component of the ATP. The three training managers took notes on the questionnaire form in order to be consistent. Training managers returned all interview questionnaires to the evaluation consultant for analysis, interpretation, and reporting.

The Report of Findings From the SCM

The training director determined the evaluation persons to receive reports. The findings were communicated using a final report, summary reports, and oral presentations. Company executive officers, general managers, training managers, and the training director received all reports, written and oral. Trainees and their regional managers received summary reports from their training managers at division meetings shortly after Seminar 1 and before Seminar 2.

In the written report, a "Summary of Interview Comments" was presented showing quotations and frequencies by categories. Training managers were also given instructions about analyzing interview information for future reference. Tables showing the comparisons between categories and content topics were presented.
The Information Analysis

A content analysis of the interview data gathered from the nine success cases studied produced data that was interpreted into frequency tables. The researcher searched the data for recurring trends and patterns to establish categories based on the number of incidents in each particular category. The system for classifying incidents evolved into six distinct categories. These classifications continued to recur when verified by a second and third review of the data. The purpose for quantifying this descriptive information was to document when, where, and how frequently the trainees were using new training skills and how it benefited them. The content analysis in Table 1 shows how the success case information was classified into the following categories:

1. Ways that trainees used their new skills.
2. How trainees had already benefited from training.
3. Changes that had taken place since training.
4. Ideas trainees had passed on to office staff.
5. Incidents when no change resulted after trying new ideas.
6. Things trainees planned to try out in the near future.

The column of total codes indicates that trainers are making more changes in certain content modules than others; therefore, it may be important to examine the data from the content analysis to consider revisions and changes for future training.
Table 1 shows an analysis of the categories of behavior coded from the success case interviews. Some specific examples of information gathered from the interviews are illustrated here:

Category 1--The number of times trainees have used their supervisory training skills to date (total 117). Examples: "I listen more intently now." "I'm using my priority list daily." "I stand up when someone walks into my office when I'm busy—this really works."

Category 2--The number of benefits the trainees could attribute to training (total 76). Examples: "I'm willing to take a risk now." "I feel more confident and more creative."
Category 3--The number of times a change could be observed (total 47). Examples: "I'm more aware of the overall picture now." "I'm more comfortable about making changes now." "I'm more aware of needs of my staff."

Category 4--The number of times the trainees passed the new ideas on to staff or friends (total 25). Examples: "I'm delegating more and my staff is more eager." "Staff is using the idea of identifying their 'Time Wasters.'" "I passed the idea on to my regional accounting office."

Category 5--The number of times no change took place or the trainees did this before training (total 19). Examples: "It's hard to pinpoint what is new and what I did before." "I need more time to review what I've learned."

Category 6--The number of times the trainees responded that they were trying to change or they planned to use the skill soon (total 18). Examples: "I haven't had time to identify my 'Time Wasters' yet." "I've practiced what I learned and I'm trying to change."

These reports were presented to several audiences within the company at local and regional meetings. Reactions and recommendations of training managers who delivered the training program and conducted the interviews were elicited and the information received from the success case interviews was used to make decisions and revisions for future training programs. A pattern of response that emerged when interviewing training managers was the need to make the problem-solving module more company specific. A final evaluation report was presented to company executive officers about 6 weeks
after completion of training. Appendix G is the Evaluation Highlights report that was written for use by general managers and regional managers, the persons immediately superordinate to the trainees. Training managers made revisions for future training programs on the basis of information received from interview data as well as data from other evaluation sources. Training materials were revised and modified in the problem solving module to make it more company-specific. Examples were taken from the success case interviews to illustrate how other trainees had used their training skills.

Meta-evaluation

The Problems Encountered in Implementing the SCM

In order to determine the problems encountered, first a group interview was conducted with the three training managers. Later the standards were selected and judgments made. The interview focused on certain key evaluative questions about how training managers used the success case selection procedures, criteria for selection, information about improving training, etc. Information received by oral and written feedback from the training managers appears in Appendix H. The results of this interview are discussed in Chapter IV.

The Strengths and Limitations of the SCM

The strengths and limitations of the SCM were determined by the researcher based on discussions with trainees, with company training
managers, and with their immediate superiors. Inferences regarding the strengths and limitations as well as the problems encountered in implementing the SCM are discussed in Table 2, Chapter IV.

Background of the Meta-evaluation

The second step of the meta-evaluation assessed the utility, feasibility, propriety, and accuracy of the SCM. The Joint Committee Standards were applied by the researcher in order to answer certain evaluation questions about the SCM. These standards are intended to assist evaluators to identify good evaluation practice and to provide formative information for decision makers regarding revisions and changes.

The Meta-evaluation Questions

The questions which guided this meta-evaluation came from the Joint Committee Standards' four major categories. These categories are:

1. Was the information provided by the SCM useful?
2. Was it feasible to implement the SCM?
3. Was the SCM implemented in a proper manner?
4. Did the SCM produce accurate information?

There were subquestions about the SCM relative to these standards which are addressed in Chapter IV, "The Findings."
The Rationale for Using Selected Standards for Making Judgments About the SCM

The researcher's selection of 10 standards from a total of 30 standards was based upon the following considerations:

1. The particular purpose of the SCM as distinguished from other evaluation approaches.

2. The four major categories of standards were considered and represented in the 10 standards chosen.

3. The information available regarding the uses and users of the SCM information.

4. The limitations as to what could be assessed about the SCM data and their use after the initial evaluation was conducted.

Based on these considerations the following standards were judged to be most salient for assessing the SCM.

A. Utility Standards: Was the information provided by the SCM useful? The utility standards are intended to ensure that the SCM served the practical information needs of the evaluation audiences.

   Audience Identification (a1). Have the evaluation audiences of the SCM information been clearly identified?

   Information Scope and Selection (a3). Have the evaluation audience needs been addressed?

   Report Dissemination (a6). Were evaluation findings disseminated to evaluation audiences?

   Evaluation Impact (a8). What follow-through was conducted by evaluation audiences to change or improve training activities?

B. Feasibility Standards: Was it feasible to implement the SCM? The feasibility standards are intended to ensure that the SCM
is realistic, prudent, diplomatic, and frugal.

Practical Procedures (b1). Did the SCM gather the necessary information in a practical way so that disruption was kept to a minimum?

Cost Effectiveness (b3). Was the SCM information of sufficient value to justify the time and resources expended to conduct the telephone interviews?

C. Propriety Standards: Was the SCM implemented in a proper manner? The propriety standards are intended to ensure that the SCM was conducted legally, ethically, and with due regard for the persons involved in the evaluation.

Rights of Human Subjects (c5). Did the nine success cases interviewed volunteer to participate in the follow-up evaluation? Was the information reported anonymously to show respect for the rights of human subjects. Was the reporting balanced? Did it address both strengths and weaknesses of the SCM?

D. Accuracy Standards: Did the SCM provide accurate information for evaluation audiences? The accuracy standards are intended to ensure that the SCM reports reveal and convey technically accurate information.

Valid Measurement (d5). Did evaluation audiences believe that the SCM instruments and procedures were developed and implemented in ways that information obtained was valid for making decisions?

Reliable Measurement (d6). Did evaluation audiences believe that the SCM instruments and procedures were developed and implemented in ways that information obtained was reliable for making decisions?

Analysis of Qualitative Information (d9). Was the SCM information appropriate and systematically analyzed to justify interpretations and recommendations?
Meta-evaluation Data Collection

Data for the meta-evaluation of the SCM were gathered by the researcher after completion of the training evaluation reports and presentations. Three success cases were selected by the researcher 2 weeks after the initial success case interviews with company training managers. One success case was interviewed from each division. Success cases were asked if the information recorded on the interview questionnaire was an accurate representation of trainers' perceptions. They were asked if the picture presented had changed with time. Results of these interviews are represented in Chapter IV.

Company training managers were surveyed about the ease of implementation of the SCM questionnaire and how costly it was to implement. The results of these interviews appear in Appendix G, "Group Interview of Training Managers." Discussion of costs, time invested, and other expenses are discussed fully in Chapter IV.

Analysis of Meta-evaluation Data

Information gathered about the SCM was used to determine compliance with the Joint Committee Standards. Selected standards were used to systematically evaluate the SCM for use by other evaluators and HRD practitioners. Conclusions about the strengths and limitations of the SCM were based upon user/client reactions and the 10 standards selected most appropriate. Conclusions and recommendations about the SCM are summarized in Chapter IV.
Recommendations to Users of the SCM

Recommendations to users of the SCM are based upon the strengths and limitations of the SCM as operationalized in the context of a particular administrator training program. The strengths and limitations are also addressed in Research Question 2 followed by a discussion of the problems encountered in implementing the SCM. Chapter V has a full discussion of these recommendations.

Summary of the Chapter

Chapter III is a summary of the design and methodology used to investigate the problem identified in this study. The success case method was operationalized, tested, and evaluated in the context of an administrator training program. The SCM was used to conduct follow-up interviews with trainees after which judgments were made regarding the meta-evaluation of the SCM. The Joint Committee (1981) standards were used to assess the SCM.
CHAPTER IV

RESULTS OF THE STUDY

Chapter IV is a summary of the meta-evaluation of the success case method. Research Question 1 described how the SCM was operationalized and tested in the context of this study. Research Questions 2 and 3 are discussed in this chapter. Research Question 4 is a discussion of recommendations for use of the SCM.

Research Question 2 states, "What problems were encountered in implementing the success case method?" The sources for this information were (a) the interview with the training managers who implemented the SCM, (b) a sample of three trainees who were reinterviewed by the researcher, and (c) the notes and reports maintained by the researcher. These sources of information provided insights into the strengths and limitations of this application of the SCM.

Sources of Meta-information

One source of information about how the SCM was used for selected trainees was the group interview conducted by the researcher. This interview with company training managers took place 1 month after training and 2 weeks after the success case interviews were conducted. The "Training Managers' Interview Questions" (Appendix H) are an indication of the feasibility and practicality of procedures and instruments to assist in assessing the SCM.
A summary of these comments shows that training managers used the suggested selection criteria in 8 of 9 cases. One person chose to select their success cases on the basis of other criteria. That person asked to change that choice to another case study after an initial interview. The training managers interviewed 8 of 9 cases by telephone and those interviewed took 30-40 minutes each. Suggestions for change included extending the time for interview and conducting face-to-face interviews, although they qualified these statements by saying, "I had to use the telephone because of constraints."

Changes that may not have shown up in other evaluation methods were the changes in attitudes of office staff and specific examples of training use which could be expanded upon during the interviews. The training managers were not yet clear on how interviews showed that they could modify or improve training. They said they were not clear or needed to delve further into how to use the information. In a later question (Question 11) the training managers responded to how trainers were using training by saying, "Many specific examples showed up, note the changes in attitudes, see how high the motivation is."

The questions asked of training managers and their responses are discussed relative to the research question about problems encountered.

Question 1: Did this method (SCM) measure changes in job performance? Responses: "Yes--trainees noticed a change in staff morale." "Attitudes in the office improved." "This follow-up encouraged trainees to discuss improvements they've seen."
Question 2: Did evaluation information document and accurately describe what had happened since training? Responses: "Yes." "Yes." "Yes."

Question 3: Did all the case studies volunteer willingly? Responses: "Yes." "Yes." "Yes." "They were all very anxious to help."

Question 4: In what ways could this method (SCM) be improved upon? Responses: "It seems easy now." "I would have preferred a face-to-face interview but had to resort to telephone because of time constraints." "Should take place later than 2 weeks after training—it's too soon to tell." "Trainees would have liked to see the questions before the interview." "I would modify to shorten the questionnaire."

Question 5: How costly was it for you to carry out the SCM? Responses: "My time to telephone three trainees which averaged about 30-35 minutes each." "The cost of the long-distance calls was far cheaper than visiting their offices." "It took a minimum of my time to present the evaluation findings to the group of trainees next time we met." "Data analysis was already summarized for my group presentation."

Question 6: Was the time spent worth the effort? Responses: "The interview was somewhat lengthy." "There was some redundancy." "Time was 30-40 minutes on the telephone which is still worthwhile." "It gave me some very specific ideas on how to improve the training for the next group." "Several examples of how to improve the problem-solving module surfaced here."
Question 7: Did any common themes occur? Were there negative consequences? Responses: "Negative consequences were that others might find out about follow-up evaluation and question why they had not been picked."

The question, "What criteria did you use to decide if you were using the training correctly or incorrectly" (Brinkerhoff, 1983, p. 58) appeared to be irrelevant in this test of the success case method. The training managers' interviews did not address this specific issue; therefore, no data were collected.

A second source of information relative to the meta-evaluation of the SCM was three of the nine success cases who were reinterviewed by the researcher. The three cases were interviewed individually a few weeks after their initial interviews, and the following questions were asked in order to assemble a profile on success case interviewees. Those interviews began with a statement about what the training managers' interviews had uncovered during the initial interview. A sample profile is shown in Figure 3.

Sample Profile on ____________________ (Name)__________________
(2 weeks after her interview)

This is the picture I got when I read what your training manager wrote about how you are using your supervisory skills training.

New skills you have used since training:

1. Plan . . .

2. Review with staff . . .

3. Identify . . .

Figure 3. Sample Profile.
After discussing the profile with the success case, the following questions were pursued.

What I need to know is this.

1. Could you answer her questions with ease?

2. Did she ask questions that showed the results of your training?

3. Is this still an accurate picture or is it distorted?

4. Did talking with your training manager help you with the questions your regional manager asked you later?

5. Did the interview seem too long, too short, or just right?

The results from these three interviews can be summarized by the following statements. Success case trainees who were reinterviewed stated that it was not difficult to respond to most of the questions asked. In a few instances they would have preferred to see the questions in advance in order to better prepare and document changes. The interview questions appeared to relate to use of new training skills in all cases. The picture as it was depicted by training managers was still an accurate representation of practice. Success cases were comfortable answering questions and did not feel threatened. The interviews that went over 30 minutes seemed too lengthy. A 20 to 30 minute interview seemed appropriate. By having to search their memories for evidence of training usage, the trainees felt they were prepared for their posttraining interviews with their superiors. This was a positive side benefit.

The interview method allowed trainees to express concerns as well as accomplishments. They especially liked communicating with
home office about changes they were making.

This method appears to be a positive approach to increasing the upward flow of information to home office. Trainees and their staffs seemed to suddenly feel involved and have new importance in the company.

**Interpretation**

The problems encountered in implementing the SCM appeared minimal to the training managers. They stated that the method was easy to use, was quick, and provided worthwhile follow-up information about trainees. One problem with telephone interviews was the lack of benefits derived from a face-to-face survey with trainees. Because trainees came from all regions of the United States and Canada, it was not feasible to interview them on the job. Training managers indicated that 2 weeks after completion of training may be too soon to tell whether success cases were actually using their new skills on the job.

The SCM provided evidence that training makes a difference. Trainers had gathered specific suggestions on how to improve certain aspects of the training. The results of these interviews helped to redesign and modify certain training modules for later groups of trainees.

**Research Question 3:** To what extent was the success case method effective and practical as an evaluation method? Meta-evaluation of the success case method is discussed relative to the four major categories of standards, i.e., utility standards, feasibility
standards, propriety standards, and accuracy standards.

The researcher's rationale for selection of 10 standards taken from the four major categories appears in Chapter III. Briefly, the purpose for using the SCM was formative and differed from other evaluation methods. The extent to which information was useful and practical to the users of the SCM information determined the standard most salient for assessing the SCM. The four meta-questions from page 52 and the standards within the four domains are as follows.

**Question 1: Was the Information Provided by the SCM Useful?**

**A. Utility Standards**

**A1, Audience identification.** The evaluation audiences for the SCM reports have been clearly identified in this study as the trainees, the training managers, the regional managers, and the company executives. The training managers were identified by company executives as the audience to present reports back to trainees and to redesign the training program based on evaluation information. The regional managers were the immediate supervisors over the trainees and needed information about changes on the job. Company executives needed the reports to make decisions regarding future funding of training efforts.

**A3, Information scope and selection.** Information collected was responsive to the needs and interests of specific audiences. Training managers received all evaluation reports with instruction on how to use information for all trainees. Training managers also had
concerns different from other audiences; they needed specific evidence to justify changes and refinements in the training curriculum. Training modifications were based on examples given to training managers by trainees. The company managers requested other equally important information in terms of effectiveness, impact on the company, and costs. These long-term issues were not specifically addressed by use of the SCM. Information selection versus information scope was considered as a way to balance the needs of the identified audiences.

A6, Evaluation dissemination. Dissemination of evaluation findings to clients was completed within 4 weeks. The audience with the greatest degree of responsibility for, and interest in, the evaluation of training was the training managers. Because they were responsible for disseminating information to the trainees, as well as for making revisions, they viewed the most comprehensive information reported. A variety of methods such as executive highlight summaries, printed reports, audio-visual presentations, and conversations were used to make information timely for each audience.

A8, Evaluation impact. Several different audiences used the SCM reports. The training managers used the evaluation information to modify content modules, especially the problem-solving module which was made more company specific because of examples gathered during follow-up interviews. Specific comments from trainees for improving the training program were, "The content should be geared more toward the individual as a manager. Too much emphasis was placed on the
problem-solving segment. It could have been condensed." "Use survival manual in discussions and outlines." "Less reading in class, insist we do our homework before the session." "Less in-class reading, more homework at night."

A second audience to receive the evaluation "Highlights" summary was the regional managers. The managers of the trainees made use of the evaluation information to document training results back on the job. For instance, managers needed to know what training skills were being used so they could observe those changes trainees said they had made. Managers said they sat in staff meetings to observe if trainees were more goal oriented, better delegators, etc. General managers reviewed evaluation results to determine future training efforts.

Question 2: Was it Feasible to Implement the Success Case Method?

B. Feasibility Standards

Bl, Practical procedures. Disruption to trainees during their work schedule was minimal. Training managers scheduled an appointment and specific times to conduct the telephone interviews. The interview took an average time of 20–30 minutes and respondents were informed of the purpose and method for collecting information. The procedures for carrying out the SCM were chosen with consideration for the overall evaluation design of this training program. The training managers considered a mail survey and agreed upon the practicality of the selection procedure and sampling plan. The
instrument was reviewed and modified to be most efficient. The time necessary to conduct the telephone interview was reasonable and more cost effective than on-site interviews. The steps necessary to implement the SCM and the analysis of data were the responsibilities of the researcher. Implementation of the method was practical and yielded usable results. Data analysis techniques presented a problem to the researcher. A content analysis was used to code perceived changes as reported by trainers to their training managers. Although the interviews were reviewed twice by the researcher and coding was found to be consistent, reporting interview data is a series of judgments. Transcripts of critical incidents when training seemed particularly applicable can be very lengthy and time consuming. To summarize all incidents in a meaningful way to various audiences is a challenge. One caveat to consider is to weigh practicality against effluence. The data must be valid and reliable to serve the practical needs of audiences.

B3, Cost effectiveness. The evaluation should produce information of sufficient value to justify the resources expended. Costs refer to the value of all resources used in the evaluation, including the time to interview trainees; the time of the interviewers; telephone time; and the time of the researcher to develop the instrument, analyze data, and write reports. Indirect costs include secretarial services, copy services, and office space. If benefits equal or exceed the costs, the evaluation is cost-effective. Benefits may include, but are not limited to, improved leadership on the part of
trainers, improved morale, improved communication between levels of employees, and improved training curriculum.

It is not always easy to adapt training costs and benefits into tangible numbers. The researcher can only attempt to quantify some of these costs as follows:

Training managers' time to report evaluation findings to trainees (3 hours @ $30/hour).

Researcher's time to train the training managers in the evaluation procedures (3 hours @ $30/hour).

Researcher's time for development and analysis of instrument (10 days @ $80/day).

Training manager's time to conduct interviews with 9 success cases (1.5 hours @ $30/hour).

Trainee's time to respond to interview questions (1.5 hours @ $20/hour).

Telephone calls to 9 trainers ($300).

Secretarial assistance and copy (5 days @ $40/day).

Estimated total costs for carrying out the SCM would be approximately $1,400. This represents about 2% of the $70,000 budget to develop, implement, and evaluate the administrator training program.

Benefits derived from the SCM included: (a) changes to improve the training content and (b) small percentage of total costs to implement (2%).

If the benefits are considered more valuable than the costs, then this method would be considered cost-effective. Training managers, when asked the question, "How costly do you consider this method?" responded, "The cost is minimal compared to traveling . . .
on-site." The training managers had considered a company-wide survey and that was not a method of choice.

**Question 3: Was the SCM Implemented in an Ethical Manner?**

**C. Propriety Standards**

**C5, Rights of human subjects.** The nine success cases all volunteered to participate in the follow-up evaluation studies. The rights of subjects were protected by reporting all data anonymously. Individuals were informed before the interview began that, "This information was not to evaluate him/her personally, but rather to improve and revise the training content and methods for the next training group." (See Appendix F, "Training Managers' Interview Form.")

On-site interviews to observe trainees may have presented a potential threat if training managers used negative changes to report to the company executives. That was not a problem in this instance.

The training managers scheduled the interviews and discussed the purpose of the interview with each trainee. The interview forms were used to construct a detailed picture relating to the SCM questions. These written instructional procedures were used by all training managers and evaluation findings were reported as a group. Other parties would not use the information for purposes different from those agreed to by the trainees.

**C7, Balanced reporting.** The evaluation was complete and fair in the presentation of strengths and weaknesses of the training
activities. This means being complete and fair in assessing and reporting both negative and positive aspects of the evaluation. One of the strengths most often reported was the amount of interaction which took place during training and as a result of training. Trainees said their staffs were accepting more responsibility in the decision-making process and they, as administrators, had more confidence to delegate responsibility. One caveat to be aware of is the tendency to manipulate the balance of strengths and weaknesses to please certain interest groups or to delete from the report weaknesses which might prove embarrassing. Within the limits of time and resources, the SCM reported the perceived comments from knowledgeable persons to represent a balance of strengths and weaknesses.

The information received showed training managers were able to modify and change certain content modules such as the problem-solving module. Several specific examples of trainees' office problems were integrated to make the content more company specific.

**Question 4: Did the SCM Provide Accurate Information for Evaluation Uses?**

**D. Accuracy Standards**

**D5, Valid measurement.** Validity is the assurance that the interpretation of the SCM data are accurate. SCM instruments and procedures were developed by the researcher after piloting and making revisions. The training managers assisted in making modifications in specific content areas. The validation process included detailed training procedures for administering the Training Managers'
Interview Form (see Appendix F). The researcher conducted a content analysis and interpreted the results to several evaluation audiences. The frequency distribution of those results appears in Chapter III.

D6, Reliable measurement. Reliability is the degree to which the SCM accurately measured what it claims to have measured. The reliability of the SCM training managers' interview form was confirmed by other evaluation information sources such as reactions of trainees taken before, during, and after training. Follow-up interview information tended to corroborate the other evaluation information sources. The researcher reinterviewed three cases to test reliability.

D9, Analysis of qualitative information. The SCM information was appropriate and systematic to support interpretations and recommendations. Revisions to the training program were based upon SCM information which provided additional depth and perception.

In order to summarize the findings based upon the interviews with training managers, the interviews with three success cases and the notes and reports provided by the researcher, a table of strengths and limitations was developed. Table 2 represents some highlights considered important to the findings. These findings are further reported in Chapter V which includes conclusions and recommendations.
## Table 2
Summary of the Strengths and Limitations of the SCM

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SCM was most useful to training managers closest to the evaluation. Evidence: Training managers in their interviews found the information helpful and useful. They used the examples of trainees to improve the training for future users.</td>
<td>1. SCM should be tested in other training situations for validity and reliability.</td>
</tr>
<tr>
<td>2. Trainers used this opportunity to interact with trainees and to reinforce learning skills. Evidence: Trainers often asked success cases for examples of how change had taken place.</td>
<td>2. Selection not random and could have built-in biases.</td>
</tr>
<tr>
<td>3. Structured interview for interviewers made questions easier to analyze for content categories.</td>
<td>3. Questionnaire format may be too restrictive and too lengthy.</td>
</tr>
<tr>
<td>4. Provided employee perceptions of training utility.</td>
<td>4. Interpretation of data was difficult to keep objective.</td>
</tr>
<tr>
<td>5. Reports were kept anonymous.</td>
<td>5. Content analysis depended on researcher's choice of categories based on expert judgment.</td>
</tr>
<tr>
<td>6. Unintended findings appeared. Specific examples were given. Evidence: Morale improved in the office. Attitudes changed for office staff.</td>
<td>6. Possibly trainees responded more positively than they would have in a written questionnaire that was anonymous.</td>
</tr>
<tr>
<td>7. Improved communications between the field and home office. Evidence: One trainer selected a case she had a previous problem with and she wanted to improve her relations with.</td>
<td>7. Self-reporting by trainees may have potential for bias.</td>
</tr>
<tr>
<td>8. Interviews conducted by trainers who delivered training may create unbalanced reporting.</td>
<td>9. The multiple rules played by the researcher were those of evaluation consultant to the project and meta-evaluator.</td>
</tr>
</tbody>
</table>
Table 2—Continued

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Cost was low versus air travel to job sites. Evidence: Telephone calls to trainees were economical.</td>
<td></td>
</tr>
<tr>
<td>9. Training managers used examples to modify training materials to be more company specific. Evidence: One training manager revised the problem-solving module based on examples of how trainees applied their new skills.</td>
<td></td>
</tr>
<tr>
<td>10. Helped top management understand how trainees were using new skills. Evidence: Examples were specific. Trainees cited times when they had delegated a duty to staff.</td>
<td></td>
</tr>
<tr>
<td>11. Success case data confirmed other reports of training effects.</td>
<td></td>
</tr>
<tr>
<td>12. The time to write reports to evaluation audiences took less than 2 weeks.</td>
<td></td>
</tr>
<tr>
<td>13. Trainers didn't need a strong background in evaluation theory to carry out the SCM. Evidence: Trainers had previous training in how to interview but no background in evaluation of training.</td>
<td></td>
</tr>
</tbody>
</table>
Analysis of Meta-evaluation Data

Information gathered about the SCM was used to determine compliance with the Joint Committee (1981) standards. Selected standards were used to systematically evaluate the SCM for use by other evaluators and human resource practitioners. Conclusions about the strengths and limitations of the SCM were based upon user/client reactions and the 10 standards selected most appropriate.

Summary of the Chapter

The results of the study are reported in Chapter IV. Research questions are addressed according to the sources of information. The standards selected for assessing the SCM are discussed giving examples when appropriate. The strengths and limitations of the SCM have been summarized. Conclusions and recommendations about the SCM are reported in Chapter V.
CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions about the meta-evaluation of the success case method can best be summarized as follows: first, by reviewing the current issues and problems in training evaluation that emerged in the search of the literature; second, by reporting the limitations of the study; and third, by examining the strengths and weaknesses of the success case method by assessing what problems it overcomes considering the study limitations. Last, recommendations will be made for further use of the success case method as an evaluation tool.

Summary of Evaluation Problems

The problems in training evaluation identified in the early chapters of the study can be capsulized as follows:

1. Evaluation methods currently in use are often very costly and complex.
2. Few evaluation methods look at the follow-up component of evaluation, the results, and impact.
3. Evaluation procedures must consider issues of utility and feasibility.
5. It is costly to observe trainees on the job.
6. Often, human resource trainers lack the necessary background or skills in evaluation to carry out complex evaluation approaches.

A comprehensive meta-evaluation of an evaluation method is a means for assessing the feasibility, utility, propriety, and accuracy of new methods. It should also provide guidelines and caveats to others interested in trying new methods of evaluation. A meta-evaluation should also stimulate further research and experimentation in the field of training evaluation.

Limitations of the Study

Certain limitations of the meta-evaluation of the success case method should be noted as caveats or trade-offs so as not to mislead others when considering similar applications for the success case method. The following possible limitations taken from Table 2 of the study were identified by the researcher.

1. The SCM was not well defined in the literature and therefore needed to be operationalized before the researcher could conduct either an evaluation or a meta-evaluation. This study future operationalized the SCM for future studies.

2. Because this test of the SCM relied upon the self-report method for collecting information about trainees, there is potential for participant bias. The telephone interviews gathered descriptive information about perceived training usage on the job. Interviews with superiors or subordinates, as well as activity samples or observer diaries may enhance internal validity.
3. The multiple roles of key participants in the evaluation may have introduced bias. Follow-up telephone interviews were conducted by the training managers. This may be a limiting factor since perceptions were gathered by the persons most responsible for the success of the training program. Additionally, the researcher assumed several roles in the study which could introduce bias. The researcher operationalized the SCM in the context of an administrator training program and designed and conducted the meta-evaluation. Independent interviewers may minimize threats to validity.

4. Finally, as indicated in Limitation 3, the meta-evaluation was not independent, as Stufflebeam (1974) recommended. If a panel of evaluation practitioners conducted the meta-evaluation of the SCM, there could be less potential for bias.

Nonetheless, the researcher believes this study has merit and can serve as a useful guide for others contemplating using the SCM of evaluation. Careful records of the test of the success case method and operational steps were kept and were integrated into this study. Success case interviewers were skilled in interview methods. The meta-evaluation was based upon information collected from several perspectives, i.e., trainees, trainers, and researcher. Recommendations including the strengths and the shortcomings of the SCM were reported.
Strengths of the SCM

Strengths of the SCM can be summarized as follows:

1. SCM may be most useful to the trainers who also conducted the follow-up evaluation because there was immediate utility to use in revising curriculum. Training managers used examples from the interviews to revise the curriculum to be more specific to trainees' experiences. The trainers said the benefits were greater than the risk of bias.

2. Interaction with trainees after training tended to reinforce training concepts for trainees. Although the questionnaire was structured, it allowed interviewers the opportunity to probe further. This probing gave trainees a means by which to influence future training efforts as well. It was a direct communication with the home office.

3. Unintended outcomes of the training were discovered. Morale improved and attitudes changed for office staffs when trainers delegated more responsibility.

4. Examples given by trainees provided specific ideas about how to improve future training efforts. The frequency distribution tables showed the greatest difficulty when applying training skills was in the problem-solving module.

5. The cost for implementing the SCM was economical in comparison to other methods. Implementation was not difficult. The telephone interviews were obviously more efficient than on-site visitations.
6. SCM data confirmed other sources of evaluation information. Evaluation questions at Levels 1 and 2 in the evaluation process showed an overall positive attitude about training. This was corroborated in the success case interviews.

7. SCM can be used by people who do not have a strong evaluation background. Training managers generally followed the guidelines suggested and used the selection criteria to select cases to study. They felt comfortable using the evaluation method.

8. The interview format and structure for the interview was flexible as well as specific to the needs of various audiences. The general success case questions were useful in formulating more sub-questions.

Weaknesses of SCM in This Context

Shortcomings of the SCM can be summarized as follows:

1. The SCM had not previously been operationalized or tested in an educational program. Research and experimentation in the field of training evaluation can be further developed.

2. Selection was not random but purposely selected successful cases. There may be biases associated with such cases which could be counterbalanced by purposeful selection of unsuccessful cases.

3. Interpretation of data from the interviews was difficult to categorize. The researcher had to depend on emerging patterns and to make judgments about the content of the questionnaire.

4. Although self-reporting by telephone is economical and useful, on-site observations of behavior changes may be more independent
of biases. Evaluation methods should be tested and refined so that others may make their own judgments regarding reliability and validity.

Recommendations to Users of SCM

The SCM is not a panacea for evaluation of all training programs but can supplement other evaluation information. The SCM may be useful to people with little formal training or background in evaluation as it is relatively simple and useful for collecting follow-up results after training. The SCM is an economical and practical method for conducting formative evaluations and produces evidence different from testimonials, i.e., specific examples for improving the content of future training.

The limitations of this study have been noted. Recommendations for potential users of the SCM may be appropriate. The following recommendations are suggestions to address the limitations as identified.

1. The potential for participant bias may have resulted from the multiple roles played by trainers who also did the follow-up interviews. A skilled interviewer, independent from the training, may enhance internal validity and reduce the potential for violation to the rights of the trainees interviewed. Having one evaluator conduct all the success case interviews also reduces the chance for distortions resulting from the trainers' involvement.

2. Because the SCM relied upon the self-reporting telephone method for collecting documentation, a visitation to the trainees'
offices may have produced further evidence of training results. Work activity sampling collected by the trainees' superiors would offset the tendency to report only positive events to company trainers. Cross-checking of trainees' reports with more documentation would increase credibility of success case findings.

3. The SCM produced information most useful to the training managers who modified and revised the content. The reports were not especially useful to higher levels of corporate decision makers. This audience for evaluation information was somewhat sensitive to cost benefits and the long-term impact of training dollars invested. The SCM did not purport to assess impact on the company.

4. A future test of the SCM may benefit from allowing more than 2 weeks to pass after completion of training before interviews are conducted. Research indicates that timing of follow-up interviews may be a critical element and that evaluation approximately 4 to 6 weeks after training may enhance reliability.

5. One difficulty the researcher had was the analysis of large quantities of data that had to be coded and recoded to check for internal consistency. Reducing the number of questions asked during the interview would have reduced the amount of time interviewers needed to complete their task, as well it would offset the abundance of data to be coded. Trainers also suggested during their meta-interview with the researcher they would shorten the length of the interviews. More careful structuring of questions and probes could reduce difficulty with these two problems.
6. Success cases interviewed reacted to the follow-up evaluations as a very positive benefit for them. They received additional attention from the corporate trainers and this could also be a limitation similar to the Hawthorne effect. The question arises, "Did the success cases get more successful because they were selected?" Some trainers would consider this a worthwhile purpose for conducting such interviews. It is possible that balance in reporting would result if an appropriate number of unsuccessful cases were interviewed also. One recommendation for future study would be to conduct follow-up interviews with persons identified in training "as persons least likely to be successful." Cross checking different data sources tends to reduce uncertainties and may increase credibility with higher level decision makers.

7. Another problem for the researcher was the potential for distortions resulting from the researcher's presence at the training site during the development, implementation, and evaluation stages of the training. A further involvement was the conduct of the meta-evaluation which required a special sensitivity on the part of the researcher to offset the tendency for bias. Selection of standards for assessing the SCM was based upon the researcher's interpretation of which standards applied to the SCM. The researcher based those judgments on what the SCM purported to do according to the author. Further study is necessary in how to apply the Joint Committee (1981) standards to educational training evaluation. The determination of appropriateness of fit between the standards and audience needs is an issue to be further researched.
Suggestions for Future Research

Based upon the recommendations of the researcher, there are suggestions for future research. Perhaps studies addressing the following questions, which became apparent during the present study, might contribute to the literature in both the field of training evaluation and the field of meta-evaluation.

1. What relationship exists between a responsive researcher involved in multiple roles during the evaluation and an independent interviewer?

2. Would cross checking trainees' self-reporting method with further documentation from superiors increase credibility of success case findings?

3. Is the timing of follow-up interviews a critical factor in assessing results and benefits of training?

4. Are there better methods for coding the frequency of events than the categories of critical incidents used in this study? How can narrative data be summarized accurately and efficiently?

5. Does the Hawthorne effect bias the credibility of data from only successful cases? What unexpected data may emerge from interviews with unsuccessful cases?

6. Are better guidelines for applications of standards needed to improve the meta-evaluation of training evaluation methods?
Chapter V summarizes some current problems and concerns in evaluating training programs. Strengths and limitations of this study are reported based upon several sources of information. Conclusions and recommendations are reported for further study.
Appendix A

Summary of Joint Committee Standards
Summary of the Standards for Evaluation of Educational Programs, Projects, and Materials

A Utility Standards
The utility standards are intended to ensure that an evaluation will serve the practical information needs of given audiences. These standards are:

A1 Audience Identification
Audiences involved in or affected by the evaluation should be identified, so that their needs can be addressed.

A2 Evaluator Credibility
The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that their findings achieve maximum credibility and acceptance.

A3 Information Scope and Selection
Information collected should be of such scope and selected in such ways as to address pertinent questions about the object of the evaluation and be responsive to the needs and interests of specified audiences.

A4 Valuational Interpretation
The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear.

A5 Report Clarity
The evaluation report should describe the object being evaluated and its context and the purposes, procedures, and findings of the evaluation, so that the audiences will readily understand what was done, why it was done, what information was obtained, what conclusions were drawn, and what recommendations were made.

A6 Report Dissemination
Evaluation findings should be disseminated to clients and other right-to-know audiences, so that they can assess and use the findings.

A7 Report Timeliness
Release of reports should be timely, so that audiences can best use the reported information.

A8 Evaluation Impact
Evaluations should be planned and conducted in ways that encourage follow-through by members of the audiences.
B Feasibility Standards
The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal; they are:

B1 Practical Procedures
The evaluation procedures should be practical, so that disruption is kept to a minimum and that needed information can be obtained.

B2 Political Viability
The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.

B3 Cost Effectiveness
The evaluation should produce information of sufficient value to justify the resources extended.

C Propriety Standards
The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. These standards are:

C1 Formal Obligation
Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

C2 Conflict of Interest
Conflict of interest, frequently unavoidable, should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

C3 Full and Frank Disclosure
Oral and written evaluation reports should be open, direct, and honest in their disclosure of pertinent findings, including the limitations of the evaluation.

C4 Public's Right to Know
The formal parties to an evaluation should respect and assure the public's right to know, within the limits of other related principles and statutes, such as those dealing with public safety and the right to privacy.
C5 Rights of Human Subjects
Evaluations should be designed and conducted so that the rights and welfare of the human subjects are respected and protected.

C6 Human Interactions
Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation.

C7 Balanced Reporting
The evaluation should be complete and fair in its presentation of strengths and weaknesses of the object under investigation, so that strengths can be built upon and problem areas addressed.

C8 Fiscal Responsibility
The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible.

D Accuracy Standards
The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features of the object being studied that determine its worth or merit. These standards are:

D1 Object Identification
The object of the evaluation (program, project, material) should be sufficiently examined, so that the form(s) of the object being considered in the evaluation can be clearly identified.

D2 Context Analysis
The context in which the program, project, or material exists should be examined in enough detail so that its likely influences on the object can be identified.

D3 Described Purposes and Procedures
The purposes and procedures of the evaluation should be monitored and described in enough detail so that they can be identified and assessed.

D4 Defensible Information Sources
The sources of information should be described in enough detail so that the adequacy of the information can be assessed.

D5 Valid Measurement
The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will
assure that the interpretation arrived at is valid for the
given use.

D6 **Reliable Measurement**
The information-gathering instruments and procedures should be
chosen or developed and then implemented in ways that will
assure that the information obtained is sufficiently reliable
for the intended use.

D7 **Systematic Data Control**
The data collected, processed, and reported in an evaluation
should be reviewed and corrected, so that the results of the
evaluation will not be flawed.

D8 **Analysis of Quantitative Information**
Quantitative information in an evaluation should be appropri­
ately and systematically analyzed to ensure supportable inter­
pretations.

D9 **Analysis of Qualitative Information**
Qualitative information in an evaluation should be appropri­
ately and systematically analyzed to ensure supportable inter­
pretations.

D10 **Justified Conclusions**
The conclusions reached in an evaluation should be explicitly
justified, so that the audiences can assess them.

D11 **Objective Reporting**
The evaluation procedures should provide safeguards to protect
the evaluation findings and reports against distortion by the
personal feelings and biases of any party to the evaluation.
Appendix B

ATP Seminar Agenda
Monday, March 21

8:30 a.m.   I. WELCOME AND INTRODUCTION TO ADMINISTRATOR TRAINING

9:15 a.m.   II. HEALTHCARE SERVICES
- Company Philosophy
- Company Expectations of Administrator

9:45 a.m.   III. HEALTHCARE SERVICES ORGANIZATION CHARTS
- Home Office
- Field Organization

10:45 a.m.  IV. DIVISION MEETINGS
- Division Philosophy
- Staff Roles
- Staff vs. Line Management
- Administrator Role
- Miscellaneous

1:15 p.m.   V. MANUFACTURING PLANT

3:00 p.m.   VI. TOUR OF HOME OFFICE
- Introduction of Individual Units
- Seminar Handbook
- Seminar Skills Inventory

Tuesday, March 22

8:00 a.m.- 12:00 noon
I. THE PHILOSOPHY OF MANAGEMENT
The Company believes that it can be successful while at the same time allowing employees to maintain self-esteem. Senior management asks all members of management to manage employees as individuals, allowing them to accomplish their personal goals while at the same time allowing the Company to reach its goals. The module will

Training Managers
- Atlantic Division - Room 6
- Central Division - Room 3
- Pacific Division - Room 9
demonstrate this concept through the use of the film "Me! and You."

1:00 p.m.- 4:30 p.m.  II. ROLE OF THE SUPERVISOR

Very few supervisors spend enough time reflecting on what is really expected of them and what are the standards for being a successful supervisor in their organization.

Even for supervisors who know the textbook definitions of supervision, this session provides what every supervisor should know in the context of the real world of management. It lets you put this into the realities of your organization, and relate the learning to achieving results. Participant ends session by designing an individual action plan for the next quarter.

Wednesday, March 23

8:00 a.m.- 1:00 p.m.  I. PROBLEM SOLVING/DECISION MAKING

Supervisors are faced with alternative ways of accomplishing results. This module concentrates on processes used to determine the best course of action and how to present the recommendations.

1:00 p.m.- 4:30 p.m.  II. LEADERSHIP

Now that we, as supervisors, know our objectives and have identified the problems of implementation, we must create a willing team to meet the goals of both the organization and the individual. This
session teaches those things that make for effective individual and group leadership and how to apply those techniques.

III. ADMINISTER MID-SEMINAR FEEDBACK

Thursday, March 24

8:00 a.m.- 12:00 noon  
I. COMMUNICATION SKILLS  
The important skills of listening and helping others is often lost in the paperwork of the job. The emphasis in this module is placed in two areas: communication and coaching/counseling employees.

1:00 p.m.- 4:30 p.m.  
II. THE ADMINISTRATOR AS A TRAINER  
Supervisors are accountable for training their subordinates. This module concentrates on how to orient new employees to the Company, identifies the function of training, and defines the supervisor's role.

5:30 p.m.  
KEYNOTE SPEAKER  
"The Fine Art of Balancing"

Friday, March 25

8:00 a.m.- 11:30 a.m.  
I. GOAL ATTAINMENT  
This module concentrates on the utilization of planning and organizing skills to accomplish short and long term goals.
11:30 a.m. - II. ADDENDUM
12:15 p.m.
  - Seminar Skills Inventory
  - Seminar Review
  - Task Planning Sheet
  - Back on the Job
  - Miscellaneous Announcements
12:30 p.m.  DEPARTURE
Appendix C

Evaluation Process
EVALUATION PROCESS

Training Manager

- Develop Interview Forms
- Interview Success Cases
- Discuss and Revise Training Materials

Regional Manager

- Observe Performance
- Review Task Planning Sheet
- Observe Performance

Home Office

- Develop Interview Forms
- Develop Task Planning Sheet
- Analyze Interview and Task Planning Sheet
- Present Evaluation Highlights
- Analyze and Revise Materials for Future Training
Appendix D

Demographics of Administrators
## ADMINISTRATOR, HEALTH CARE SERVICES
### Seminar I - Supervisory Skills
#### Demographic Information

37 Respondents

<table>
<thead>
<tr>
<th>1. Sex:</th>
<th>Male</th>
<th>(1)</th>
<th>Female</th>
<th>(36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Age:</td>
<td>Under 25</td>
<td>(1)</td>
<td>26-35</td>
<td>(10)</td>
</tr>
<tr>
<td></td>
<td>36-45</td>
<td>(14)</td>
<td>46-55</td>
<td>(9)</td>
</tr>
<tr>
<td></td>
<td>Over 55</td>
<td>(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Marital Status:</td>
<td>Single</td>
<td>(3)</td>
<td>Married</td>
<td>(26)</td>
</tr>
<tr>
<td></td>
<td>Separated</td>
<td>(2)</td>
<td>Divorced</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>Widowed</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Years in present office:</td>
<td>Less than 1</td>
<td>(5)</td>
<td>1-4</td>
<td>(20)</td>
</tr>
<tr>
<td></td>
<td>5-8</td>
<td>(7)</td>
<td>9-12</td>
<td>(5)</td>
</tr>
<tr>
<td>5. Years with company:</td>
<td>Less than 1</td>
<td>(2)</td>
<td>1-4</td>
<td>(17)</td>
</tr>
<tr>
<td></td>
<td>5-8</td>
<td>(12)</td>
<td>9-12</td>
<td>(6)</td>
</tr>
<tr>
<td>6. Formal Education:</td>
<td>High School</td>
<td>(5)</td>
<td>College Degree</td>
<td>(7)</td>
</tr>
<tr>
<td></td>
<td>Graduate Degree</td>
<td>(5)</td>
<td>RN</td>
<td>(14)</td>
</tr>
<tr>
<td></td>
<td>LPN</td>
<td>(2)</td>
<td>Additional Courses/Seminars</td>
<td>(10)</td>
</tr>
</tbody>
</table>
Appendix E

Demographics of Success Cases
1. Sex: Male (1)  Female (8)

2. Age: Under 25 (3)  26-25 (3)  36-45 (3)  46-55 (3)  Over 55

3. Marital Status: Single (8)  Married (1)  Separated (1)  Divorced  Widowed

4. Years in present office: Less than 1 (3)  1-4 (4)  5-8 (2)  9-12

5. Years with company: Less than 1 (1)  1-4 (5)  5-8 (3)  9-12

6. Formal Education: High School (2)  College Degree (4)  Grad. Degree (1)  RN (1)  LPN  Additional Courses/Seminars (4)
Appendix F

Training Manager's Interview Form
One task in the Evaluation Process is to measure if, and specifically how, training skills and knowledge have changed since the seminar. As a Training Manager, you can meet this objective by interviewing a few Administrators to document if learning took place. The following guidelines assist you to organize and document specific and accurate information in a useful manner.

Not only will this information be helpful to you as a trainer, but it will provide the Home Office Training Unit with data to revise and improve future training efforts.

Please review the entire form thoroughly before you schedule your interview with the trained Administrator. Try to be as descriptive and clear as possible in recording. Assure the Administrator that this information is not to evaluate him/her personally, but rather to improve and revise the training content and methods for the next training group. If you can arrange a personal interview, schedule it during an office visit. If this is not feasible or cost effective, a telephone interview should be scheduled to discuss the topics to be covered.

Make notes of the attached sheets on "Hints for Asking Questions" and "Follow-Up Questions."

PROCEDURES

Step 1. Discuss the purpose of the interview.

"My purpose in calling today is to ask questions about how you might be or might not be, using certain skills you learned at the Training Seminar. This information will assist us in making changes and revisions to improve future training, not to be judgmental about you.

"The questions I'll be asking focus on the topics covered in Seminar I, that is:

- Role of the Supervisor
- Problem Solving/Decision Making
- Leadership
- Goal Setting/Time Management
- Administrator as a Trainer
- Communications and Coaching

"I'll be taking a few notes as we talk so I don't forget important information."
Step 2. The purpose of the interviews is to construct a detailed picture relating to these questions:

- How have they used the training?
- What benefits can be attributed to the use of training?
- What problems did they encounter in using the training?
- What negative consequences ensued from the training and/or its use?
- What criteria did they use to tell them they were using the training correctly? (or incorrectly?)

Step 3. Utilize the Interview Forms given here to elicit specific answers to your questions. If you depart from the questions on the blue sheet, the general questions will help you focus direction. Upon completion, make a copy of your Interview Form and mail it to ________ in the Home Office. Ask probing questions without leading the person.

Step 4. Close the interview on a positive note. Thank the Administrator for his/her time and information and assure him/her again that this is being used to evaluate training and not his/her job performance.

ALLOW TIME FOR THEM TO ANSWER. LISTEN FIRST SO YOU CAN CONCENTRATE ON THOSE PARTS OF THE COMMENTS THAT RELATE DIRECTLY TO THE QUESTIONS ASKED.
INTERVIEW FORM

I. ROLE OF THE SUPERVISOR

1. While in Training Seminar, you learned some fundamentals to make your job easier, such as organizing, delegating, communicating. Have you changed the way you organize your work: What would be an example?

2. Did you encounter problems when you delegated a job? If so, what should have been done that wasn't?

3. Did you bring any new ideas to your job regarding your delegating skills? What was the outcome? Are you getting feedback?

4. Can you describe a situation when you helped an employee find his/her own solution to a problem?

5. Can you show me/share with me any work samples that have made a positive/negative change in your supervisory role since training? Any additional comments? Have you tried different ways of motivating your staff? Explain.

II. PROBLEM SOLVING/DECISION MAKING

6. The Problem Solving module contained some Pre-Workshop Exercises where you had to describe some business problems you would like to solve. Was it helpful to identify these problems going into the Seminar, and did it make your task easier?

7. Were you able to apply the Problem Solving steps after you got back to your office? If not, why? If so, describe a situation when you did.

8. Have you used any of the Seminar strategies to enable you to plan creative solutions? What would be an example?

III. LEADERSHIP

9. According to the Situational Leadership Theory you earned at Seminar, we analyzed task behavior and relationship behavior. Can you think of a situation when you changed your own style of leadership in the past two weeks? What was the outcome?
10. Do you feel you are more or less adaptive in your style of leader behavior since training? When?

11. Have you changed your "task" or "relationship" behavior with your staff as a result of training? Tell me more about that.

IV. GOAL SETTING/TIME MANAGEMENT

12. In the Goal Setting module we concentrated on your planning and organizing skills. Has this helped you? What benefits/problems have you encountered when you set only important goals?

13. Are there ways you can improve how you manage your time? Explain.

14. Have you been able to identify some of your "time wasters"? Tell me exactly what you do now to avoid "time wasters."

15. What do you do now that is different?

V. ADMINISTRATOR AS A TRAINER

16. In the Seminar you learned to define "training" and discussed your role in the training process. Have your perceptions changed since training? Does this make your job easier or harder? Why?

17. What new ideas can you use to determine your office staff training needs?

18. How does that compare with your efforts before training?

19. Can you show me _________________ ... or ... Can you describe the process, step by step.

VI. COMMUNICATIONS AND COACHING

20. You practiced skills of listening and speaking while in training. Have you tried "Active Listening" in the past week? How often? Did it enhance understanding?

21. What evidence of changes at work are directly related to your new communicating skills? Tell me more about that. What was the outcome?
22. Do you have additional comments to improve the training experience for the future? How else _________________?

CLOSE BY THANKING THE ADMINISTRATOR FOR HIS/HER TIME.
Appendix G

Evaluation Highlights
Attached is the Evaluation Highlights report for Seminars I and II, which reflects company-wide results of Administrator training thus far.

Each General Manager and Training Manager has the complete report, called Administrator Training Program-Evaluation Seminar I and II, which covers all levels of evaluation for the company. The Training Managers were also given all the comments and tallies for their respective Divisions.

If you would like a copy of the complete report, contact either your General Manager or your Training Manager. If you have any questions regarding the findings, please contact me.
The Training Program for Administrators consists of four seminars. As you can see by the dates, the first two seminars have been completed. The goal for each of the Training Seminars is to permit learning to take place thereby effecting change back on the job. A very important step in the learning process is the feedback of information gathered from Regional Managers, Training Managers, and Administrators. In training and development terms this is called "evaluation." Evaluation provides valuable information for making decisions regarding modifications and revisions for future training.

The evaluation model for this training program is more extensive and long-range than used in the past. The effectiveness of the evaluation process is only as useful as the information gathered. Follow-up training by Training Managers and Regional Managers determines how the value of a training program can be measured. The following Evaluation Levels explain how measurement of training took place.

**Evaluation Levels**

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
<th>Data Sources</th>
</tr>
</thead>
</table>
| Level 1--Reactions | How well did Administrators "like" the program? How did they feel? | - Mid/Post Seminar Feedback  
- Daily Reaction Sheets |
| Level 2--Learning | To what extent did they learn the skills and knowledge taught? | - Pre/Post Seminar Skills Inventory |
| Level 3--Results | To what extent did behavior change? Did they apply what they learned? | - Case Study interviews conducted by Training Managers |
| Level 4--Impact | Did training have any impact? Did it change sales, profits, morale? | - Observations/Interviews conducted by Regional Managers  
- Review of sales, earnings, and turnover statistics |

Each of these levels answers a different set of questions and each level serves a different purpose.
SEMINAR I--RESULTS

Level 1--Reactions

Overall, Administrators' reactions at Level 1 were very positive. They especially liked the interaction between divisions and they said the content was good. Complete summaries of results for seminar content, how the Administrators rated the Training Managers and what they liked "best" and "least" are available upon request. Suggestions were received to condense the materials presented and to make moderate revisions of one or two modules.

Level 2--Learning

Before and after the training seminar the Supervisory Skills Inventory was administered to all Administrators. The average score on the pre-test was 17 right out of 20 questions. At the close of the program, the average score was 18.5 correct. This is an indication that learning took place.

Level 3--Results

Case study interviews conducted by Training Managers indicate that 2-3 weeks after training Administrators are "using" their new training skills or they have passed ideas on to the office staff and others. In some cases they indicate "it's too soon to tell yet, but they plan to try to make changes." We have many experiences to document the Level 3 results of training as change is taking place back on the job.

Level 4--Impact

Interviews/observations have been conducted by the Regional Managers before and after training. Feedback from this perspective has been very meager. Regional Managers are the only persons in this evaluation process who can observe Administrators in their natural setting. The interview data has been analyzed for the few pairs of pre- to-post observations forms returned and typical responses say:

<table>
<thead>
<tr>
<th>Before Training</th>
<th>After Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;I had limited supervisory training.&quot;</td>
<td>&quot;Training has made my job easier and more rewarding.&quot;</td>
</tr>
<tr>
<td>Regional Manager observed Administrator's need to improve in setting an example for setting goals.</td>
<td>Regional Manager observed Administrator conducting a staff meeting--showed she was well organized, etc.</td>
</tr>
</tbody>
</table>
The following is a summary of trainees' comments from nine cases interviewed by three Training Managers. Each Training Manager contacted three trainees 1 to 2 weeks after Seminar I and asked a series of questions regarding how they had used their new supervisory skills and what evidence of change they could attribute especially to their training experiences. They asked 22 questions of these trainees with follow-up probes asking for consequences. In several instances, the trainees had already seen evidence of change, either in attitude or in improved efficiency for themselves and their office staffs. Many felt more confident, more positive, and more willing to take risks at being creative. Attitudes had changed for the positive and most were planning to try those things they had not had time to do yet. They wanted more time to identify changes, as this was somewhat early to make such an assessment. The key questions asked were:

1. How have you used your new supervisory training skills? Tell me exactly what you did. (No. of Codes = 117)

I'm more motivated; I'm giving better directions now. My Coordinator is doing a Direct Mailing I delegated. I listen more intently now. I'm treating my office staff differently, keeping them moving and interested. I'm using my priority list daily now. I set aside blocks of time now, and prioritize my activities.

It's still too soon to tell, no big changes yet, but I feel more positive.

2. What benefits can you attribute to the use of your training? Give me an example. (No. of Codes = 76)

Staff is using the idea of identifying "Time Wasters"--they see me doing it. I passed the idea on to my Regional Accounting Office. I'm more aware of the overall picture now, I'm willing to take a risk with my staff now, I feel more confident and more creative. I'm more comfortable about making changes now. I'm trying to let go and delegate more now. I'm forcing myself to change the way I listen. I'm planning and mapping things out now. I'm using my ideas to organize my bowling league.

3. What changes in attitude have you observed? Describe a situation. (No. of Codes = 47)

My staff is more eager, I feel more confident, feel more positive because I know where to get the answers. I was surprised to learn about my leadership style, almost refused to believe it. Now I know what I can work on. I stand up now when someone
walks into the office and I'm busy--this really works. I'm more aware of needs of my staff, the new Nursing Supervisor needs time to review all books, tapes, etc. My Regional Manager is very supportive.

4. What problems have you encountered in trying to use your training? Why do you think this happened? (No. of Codes = 28)

I did not use some of the skills yet, it's too soon to tell. I'm too busy to try some things yet. It's hard to identify how I changed because I've blended what I already do. There is still a gap between the Administrators and the Regional Managers--communications about the whole program have not changed some attitudes. I've practiced what I learned and I'm trying to change. I'm trying some new things but I'm so busy with crises... I need more time to review what I've already learned, I'm already reading new pre-seminar materials for the next session. I haven't had time to identify my "Time Wasters" yet, I will think more about it. It's hard to pinpoint what is new and what I did before.
Appendix H

Group Interview of Training Managers
Group Interview With Training Managers and Their Responses

The researcher met with the three company training managers to collect information about how useful the SCM was to aid in selection and interview of success cases studied. This group interview took place one week after completion of the success case study interviews. The following questions were generated to elicit information for the meta-evaluation of the SCM. Training managers' responses are summarized here:

1. Did you use the selection procedures and criteria you were given when choosing your success cases? If not, why?

"Yes, I used the criteria to select those who would be open and honest." "Yes, I wanted to find out where one person was coming from."

"Yes, in two out of three cases. One seemed withdrawn and I wanted to see if I could enhance our relationship by getting to know her better." (This training manager later changed this case.)

2. Did you interview your success cases in person or by telephone? How long did it take and what would you have preferred?

Eight of nine interviews took place by telephone. In all cases the interview took 30-40 minutes by telephone; the on-site interview took 1 hour. The training managers would have preferred on-site interviews but they did not have enough time or money to do so.

3. Was the information useful to you? Was the method effective in documenting application of training skills?

Yes, I got parallel information from all three cases interviewed. This confirmed what the other evaluation methods had gathered. "Maybe other selection criteria would have shown more diversity."

4. What unexpected benefits/consequences showed up during your interviews?

Many specific examples of attitude changes in members of their office staffs. One case said, "During a creative problem solving session, the staff came up with the idea of... I had misidentified the problem and they resolved it quickly." Another case said, "My staff views our newly implemented brain storming activity as something very positive." When cases were asked if
it was helpful to identify some business problems they would like to work on, their responses indicate: Yes, it was definitely helpful, it's making resolution a little easier; yes, it was most valuable, it forced me to write it down (we usually don't do that). One case studied said, "I did not list organizational problems because they were out of her jurisdiction, it was up to her boss, and she could do nothing about the problems."

5. Was the interview questionnaire useful, too lengthy, too complex? What problems did you have using it?

One said it was too lengthy and some redundancy. It was still worthwhile and they would use it again. They would modify slightly.

6. Did the interview information show how you could improve or modify your training?

"Yes." "Not clear yet, needs further delving." They later stated they would use specific examples from the interviews to make the problem solving module more company specific.

7. Next time would you interview fewer or all your trainees?

They all said they would interview all trainees, they didn't want some persons to feel left out.

8. Will you interview the same three success cases next time you have a training seminar or will you change your cases?

Two of three training managers will interview the same cases they identified as success cases. One training manager wants to change one case and use someone considered successful (no further explanation).

9. Did you make any discoveries about training applications that may not have been obvious using other methods?

Many specific examples showed attitude changes in their office staffs. Motivation increased dramatically and this was a surprise to training managers.

10. Did any patterns or themes recur?

"Others might find out they had not been selected for follow-up and this would be a negative consequence."
BIBLIOGRAPHY


