4-1988


Sarah P. Gutek
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations
Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/2183

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.
TRAINING-PROGRAM EVALUATION: AN INVESTIGATION OF PERCEPTIONS AND PRACTICES IN NONMANUFACTURING BUSINESS ORGANIZATIONS

by

Sarah P. Gutek

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 April 1988
Despite the proliferation of evaluation literature and business and industry's growing investment in training activities, the evaluation component of training programs remains underdeveloped. Without evaluating the merit and worth of training, there is little evidence that organizations are getting a return on investment from training programs they sponsor.

The purpose of this study was to investigate the discrepancy between the emphasis on training-program evaluation in professional literature and actual practice in the field. The writer examined the perceptions of two stakeholders in the training process: upper-level managers and training directors.

Three hypotheses concerned the relationship between the frequency with which training evaluations were carried out (dependent variable) and the congruence in upper-level managers' and training directors' perceptions regarding the (a) value of evaluation activities in generating management support, (b) need for evaluation activities in helping produce better training, and (c) feasibility of conducting training evaluations in their organizations (independent variable). Three additional hypotheses
investigated whose positive agreement on the value of, need for, or feasibility of training-evaluation activities (independent variable) was related to an actual increase in the frequency of training-evaluation activity (dependent variable).

Data were collected from a random sample of 120 nonmanufacturing businesses with over 1,000 employees. Chi-square, Cramer’s V, and lambda were used to measure perceptions on value, need, feasibility, and degree to which training-evaluation activities were generally carried out. I-tests were performed to discover differences in evaluation activity when upper-level managers and training directors disagreed on the value of, need for, or feasibility of training-evaluation activities.

Major findings of the study were:

1. Training remains a corporate exercise that is taken on faith, with little or no demand to evaluate it rigorously.

2. The strongest relationship existed between perceived feasibility of the training-evaluation activity and the frequency with which the evaluation occurred.

3. The training director’s positive perception of the value of, need for, or feasibility of an evaluation activity had the most influence on frequency of the evaluation activity.

4. Where there was positive agreement concerning evaluation activities, there was increased frequency of training evaluation.
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Training-program evaluation: An investigation of perceptions and practices in nonmanufacturing business organizations

Gutek, Sarah P., Ed.D.
Western Michigan University, 1988
ACKNOWLEDGEMENTS

Although writing a dissertation often seemed like an impossible task, perhaps as Lewis Carroll’s Alice found out, believing the impossible just takes practice. Many people helped this effort by their believing. Among those I wish to thank are:

Dr. Kenneth Dickie, chairperson of my committee, who provided faith, structure, and support all the way.

Others on my committee: Dr. Dale Brethower, Dr. Robert Brinkerhoff, and Dr. Dave Cowden for their valued suggestions and assistance;

Ed Delana, statistical consultant, for his patience and time in assisting with the statistical procedures and analysis;

My mother, Jacoba Dahm, whose help in tasks large and small was a valued source of encouragement;

My immediate family, Rob, Shelly, Suzanne, and Chele, for their tolerance, love, support, and strategic use of humor;

My professional family and co-workers, what would I have done without each of you!!

Sarah P. Gutek
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................. ii  
LIST OF TABLES ...................................................................................... vii  

CHAPTER  

I. INTRODUCTION .......................................................................... 1  
Importance of the Study ............................................... 1  
Purpose of the Study ....................................................... 14  
Organization of the Study ........................................... 15  

II. REVIEW OF RELATED LITERATURE ........................................... 17  
Training-Evaluation Models ........................................... 17  
Need for, Value of, and Feasibility of Training Evaluation ....................................................... 32  
Need for Training Evaluation ................................ 33  
Value of Training Evaluation ................................ 35  
Feasibility of Training Evaluation .................... 38  
Research Questions ........................................................... 43  
Hypotheses Tested ........................................................... 43  
Assumptions ...................................................................... 44  

III. METHODOLOGY AND PROCEDURES ............................................... 50  
Development of the Instrument .................................... 50  
Sampling Procedure ........................................................... 54  
Data Collection .............................................................. 55  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents--Continued

Treatment of the Data ....................................................... 57
Analysis Procedures ............................................................ 57

IV. ANALYSIS AND DISCUSSION OF RESULTS ....................... 61
Introduction ................................................................. 61
Questionnaire Return .......................................................... 61
Characteristics of the Population ........................................... 62
  Type of Organization .................................................. 63
  Number of Employees in the Organization ......................... 64
  Number of Employees Participating in Training .................. 64
  The Form Training Takes in the Organization ..................... 66
  Importance of Training to the Organization's Success ........... 66
Characteristics of Evaluation Activities, as Reported by Training Directors ........................................ 68
  What Kind of Training Is Evaluated and How Often? .............. 68
  Satisfaction With Training Evaluation ............................. 69
  Motivation to Accomplish Training Evaluation .................... 72
  Who Conducts Training Evaluation ................................... 73
  With Whom the Evaluation Results Were Shared ................. 73
  Management Requiring Training Evaluations ....................... 76

Perceptions of Upper-Level Managers and Training Directors Concerning the Value of, Need for, and Feasibility of Training-Evaluation Activities and the Degree to Which the Activities Were Carried Out ........................................ 76
Table of Contents--Continued

Agreement Between Upper-Level Managers and Training Directors Concerning Evaluation Activities and the Frequency With Which the Activities Were Carried Out .................. 77

Frequency of Training-Evaluation Activity for Groups Reporting Agreement and Disagreement ... 81

The Association Between Positive Agreement, Disagreement, and Negative Agreement on the Part of Upper-Level Managers and Training Directors and the Degree to Which the Evaluation Activities Were Generally Carried Out .... 85

Disagreement Between Upper-Level Managers and Training Directors and How It Affects the Frequency of Evaluation Activities .................. 92

A Comparison of the Frequency With Which Evaluation Activities Were Accomplished Between Respondents Who Were and Those Who Were Not Satisfied With the Quality of Their Training Program Evaluations .................. 98


V. SUMMARY, CONCLUSIONS, RECOMMENDATIONS, AND IMPLICATIONS ........................................... 103

Summary .................................................. 103

Conclusions ............................................. 108

Conclusions Related to Hypothesis 1 Through 3 ........................................ 109

Conclusions Related to Hypotheses 4 Through 6 ........................................ 111

Conclusions Unrelated to the Hypotheses ........................................ 113
Table of Contents--Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendations</td>
<td>117</td>
</tr>
<tr>
<td>Implications</td>
<td>120</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A. SURVEY INSTRUMENTS</td>
<td>124</td>
</tr>
<tr>
<td>B. LETTER AND POSTCARD SENT TO COMPANY PRESIDENTS</td>
<td>132</td>
</tr>
<tr>
<td>C. COVER LETTER SENT WITH QUESTIONNAIRES</td>
<td>135</td>
</tr>
<tr>
<td>D. SECOND COVER LETTER</td>
<td>140</td>
</tr>
<tr>
<td>E. POSTCARD REMINDER</td>
<td>143</td>
</tr>
<tr>
<td>F. THIRD COVER LETTER</td>
<td>145</td>
</tr>
<tr>
<td>G. FOURTH COVER LETTER</td>
<td>147</td>
</tr>
<tr>
<td>H. PARTICIPANTS WILLING TO BE CREDITED FOR RESPONDING</td>
<td>149</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>152</td>
</tr>
<tr>
<td>Table Number</td>
<td>Title</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Distribution of Respondents by Type of Organization</td>
</tr>
<tr>
<td>2</td>
<td>Number of Employees in the Organizations</td>
</tr>
<tr>
<td>3</td>
<td>Number of Employees Participating in Training</td>
</tr>
<tr>
<td>4</td>
<td>Percentage of Training Provided by Various Groups</td>
</tr>
<tr>
<td>5</td>
<td>Importance of Training to the Organization's Success</td>
</tr>
<tr>
<td>6</td>
<td>Evaluation of Internal Training</td>
</tr>
<tr>
<td>7</td>
<td>Evaluation of Training Given by External Consultant</td>
</tr>
<tr>
<td>8</td>
<td>Evaluation of On-The-Job Training</td>
</tr>
<tr>
<td>9</td>
<td>Satisfaction With the Quality of Training Evaluation</td>
</tr>
<tr>
<td>10</td>
<td>Percentage of Respondents Reporting Particular Items as Impediments to Quality Training Evaluation</td>
</tr>
<tr>
<td>11</td>
<td>Mean Ranking of Factors Motivating Training Directors to Evaluate Training</td>
</tr>
<tr>
<td>12</td>
<td>Ranking of Motivating Factors in Evaluating Training by Those Required and Not Required to Share Results</td>
</tr>
<tr>
<td>13</td>
<td>Percentage of Training Directors Who Identified the Following as Conducting Training-Evaluation Activities</td>
</tr>
<tr>
<td>14</td>
<td>Percentage of Training Directors Who Identified Persons With Whom Training-Evaluation Results Were Shared</td>
</tr>
<tr>
<td>15</td>
<td>Frequency With Which Evaluation Activities Were Reported to Be Generally Carried Out and the Percentage of Upper-Level Managers (ULM) and Training Directors (TD) Who Agreed That the Various Activities Were Feasible, Would Produce More Effective Training, and Would Help Generate Management Support</td>
</tr>
</tbody>
</table>
16. Frequency of Training-Evaluation Activity, Reported by the Means for Three Groups: Those Reporting Positive Agreement, Those Disagreeing, and Those Reporting Negative Agreement ....................................................... 82

17. Association Between Agreement or Disagreement About the Value of the Training-Evaluation Activity to Help Produce Management Support and the Frequency With Which the Activity Is Carried Out .......................... 87

18. Association Between Agreement or Disagreement About the Need for Training-Evaluation Activity to Help Produce More Effective Training and the Frequency With Which the Activity Is Carried Out ............. 89

19. Association Between Agreement or Disagreement About the Feasibility to Carry Out the Training-Evaluation Activity in the Organization and the Frequency With Which the Activity Is Carried Out ............................................... 91

20. Mean Scores of Actual Evaluation Activity When Upper-Level Managers (ULM) and Training Directors (TD) Disagreed on Whether the Activity Would Generate Management Support, Help Produce Effective Training or Would Be Feasible in Their Organization ................................. 93

21. A Comparison of Evaluation-Activity Means Between Training Directors Who Were Satisfied With the Quality of Training-Evaluation Activity and Training Directors Who Were Not Satisfied With the Quality of Training-Evaluation Activity ................................................. 100

CHAPTER I

INTRODUCTION

Importance of the Study

The development and maintenance of highly educated and skilled human resources are of great economic importance. American society in the 20th century has been transformed from an agricultural to an industrial and finally to a postindustrial economy. This transformation has resulted in a dependence on human capital rather than on physical capital. Human capital, as defined by economists, consists of acquired energy, motivation, skills, and knowledge that can be used over a period of time to produce goods and services (Bowen, 1977).

One of the first attempts to tie the growth of real national income to the increased skill of the workforce was undertaken by Denison (1962, 1974) using a technique called the "residue effect." Denison isolated as many factors for growth due to identifiable sources as possible and eliminated them; he attributed the substantial residue (what was left) as being due to the effect of education. He concluded that education was a significant factor in American economic growth from 1920 and 1969. In 1979, Kendrick (cited in Barton, 1982) used Denison's data but measured the effect of both education and training more directly. Out of a total factor 1
productivity increase rate of 2.7% per year from 1948 to 1966, Kendrick credited education and training with 0.6%. In the period from 1966 through 1977, that figure increased to 0.7% of the much lower annual rate of 1.3%.

Recent research at Harvard University also provided evidence of the importance of education and training. In an analysis of the direct effect of human work on economic growth between 1948 and 1973, Jorgenson (cited in Carnevale, 1983) found that the total contribution of labor inputs accounted for 28% of that growth. In addition, 45% of the increase in the contribution of labor to growth was due to the quality of labor. The quality of labor depends on the education, training, and health care provided to the workforce. According to Jorgenson, economic returns to the individual and society from investments in education, on and off the job, were roughly 10 times those reported in previous economic literature.

Two economists, Ginzberg and Vojta (1981), reported the results of recent shifts in the focus of the economy. Their data showed that the provision of services has displaced the production of goods as the principal economic activity in the United States. In 1929, 45% of the working population was employed in the production of goods; by 1977 that sector had dropped to 32%. Employment in the service sector, meanwhile, increased to 68% of the working population, most of the shift occurring in the three decades after World War II.
Drucker (1969) supplied another view of this transformation. In describing the rise of the postindustrial workforce, he pointed out that industry, which was formerly experience based, is now knowledge based. The experience-based worker, developed through apprenticeships, has been replaced by the knowledge-based worker, developed through education and training. As a result, knowledge has become the central economic resource embodied in the form of human capital. Expanding on this notion, Carnevale (1983) produced data indicating that one of the principal contributions to productivity in recent years has been what he termed "working smarter"--the application of one's education, training, and on-the-job experiences to processes and production.

In addition to the shifts that are taking place, there is also the factor of accelerating technological change, with which the nation's workforce needs to keep pace. "As the pace of technological change accelerates, our ability to adapt our human skills to the new technologies and integrate new production techniques with available labor will be critical" (Carnevale, 1982, p. 110).

According to demographic data, 80% of the people who will be in the workplace in the year 2000 are already there ("Getting Ready," 1986). As the pace of technological change accelerates, it is likely that the job an employee has now is not the job he/she will have 20 years from now.
Whether the focus is on a shift from an industrial to a postindustrial society, from a production-based economy to a service economy, or from experience-based workers to knowledge-based workers dealing with accelerated change, the conclusion remains that the economic health of society depends largely on providing a high level of knowledge and skill to a large portion of the workforce. It is also obvious that knowledge and skills are not static but highly dynamic; they require continuous and systematic updating to keep pace with accelerated change (Lynton, 1984). The concept of human capital as the necessary investment in the continuous and systematic updating of the workforce’s skills and knowledge is the primary impetus for training by business and industry.

To train, educate, and develop employees, corporations alone are spending at least $30 billion annually and are involving 15 million people, or 1 in every 8 working Americans, in 17 million formal programs each year (Carnevale, 1986; Eurich, 1985). Training in business and industry has been divided by Nadler (1984) into three categories: learning experiences that (a) improve performance on the employees present job, or training; (b) prepare the employee for an identified job in the near future, or education; and (c) provide for the general growth of the employee, or development. For the purposes of this study, the writer uses the broader definition of training as all of the learning experiences provided for employees to be certain that their behavior contributes to the goals
and objectives of that business or industry (Friedman & Yarbrough, 1985; McGhee & Thayer, 1961; Tracey, 1984).

The growth in training efforts by business and industry has paralleled the change from agriculture to industrial and then to postindustrial society, with its shifting emphases and evolving patterns of employment. The size and scope of corporate training are difficult to measure and describe because training has grown and flourished behind closed doors.

John Dunlop, a former U.S. Secretary of Labor, coined the term "shadow system of education" to describe the emerging employer-sponsored instruction existing in virtual isolation and with little attention from established academic institutions (cited in Lynton, 1984). One reason for the isolation and lack of attention is that training programs are frequently decentralized, operating in company branches or offices scattered across the country. The development of training programs is often pragmatic; programs arise as the need arises and often disappear as quickly (Eurich, 1985).

In 1977, Lusterman revealed the full extent of employer-sponsored education with the first quantitative study on the subject. His report, entitled Education in Industry, left little doubt about the magnitude of the "shadow system of education." The study was based on a survey of 610 corporations, all with 500 or more employees, and combined the research techniques of questionnaire and personal interview. The results showed that, in 1976, 70% of all companies provided some in-house courses for their
employees, 89% had tuition-aid or refund programs, and 74% authorized outside courses during working hours.

In a 1985 report entitled Trends in Corporate Education and Training, Lusterman focused on the kinds of changes that major firms had initiated since 1977 and the factors that had prompted those changes. The report was based on information provided by 218 companies. Lusterman noted that rapid technological changes and the adoption of new strategies in response to global competition, deregulation, and other changes in the business climate all had an effect on training. A larger percentage of employers in all major job categories were involved in formal training than in 1977. In addition, Lusterman reported that the number of professionals providing training had increased in about half of the surveyed companies. Growth in both number of trainees and training staff size was particularly strong among financial institutions.

A recent report released by the Bureau of National Affairs (1985) included figures based on a cross-section of 159 corporations, 50% of which had more than 1,000 employees and 50% of which had fewer than 1,000 employees. Nine out of 10 companies surveyed conducted formal training programs for management and nonmanagement employees. Tuition assistance for college-level courses was available in about three fourths of the companies surveyed, and nearly as many companies would cover the cost of attending outside seminars or workshops.
A comparison of the 1977 Lusterman study and the 1985 Bureau of National Affairs study indicated that the number of in-house courses has grown significantly in the past 8 years. Further measures of the size and growth of the system are as follows:

1. Drucker (1973) observed that in 1940 only one corporation in the United States and one in Great Britain conducted management programs. In 1950, more than 3,000 corporations were conducting such programs.

2. A U.S. Training Census and Trends report that was summarized in the October 1983 issue of Training showed a total of 213,000 full-time trainers in the United States in 1983, with an additional 786,000 part-time trainers (Zemke, 1983). In a more recent study published by the Conference Board, it was found that in 1986 the number of trainers had increased in nearly two thirds of the firms surveyed (American Society for Training and Development [ASTD], 1986).

3. By the end of 1985, American organizations collectively delivered approximately 1,377 billion hours of formal training. Each hour of training means one employee trained for an hour; formal means the training is structured and not of the on-the-job variety (Feuer, 1985c).

4. Membership in the ASTD doubled from 1974 through 1980 to more than 21,000 national members with a possible 20,000 holding only local memberships in one of the 127 ASTD chapters throughout
the United States (Craig & Evers, 1981). By 1984, ASTD had grown to about 50,000 members nationwide (Lynton, 1984).

One can make only very rough estimates of the total size of training in terms of costs. Estimates may vary widely, depending on what is included in that estimate. Two major questions must be considered. First, does one count the trainee's wages and benefits during the training period? Second, does the cost of training reflect on-the-job training? When training expenses are figured on instructor, course-development, facility, and tuition costs, they are referred to as formal employee training costs. One estimate made for ASTD in 1981 set formal employee training costs at about $30 billion annually (ASTD, 1986). Several other studies have suggested that training costs would at least double if figures for employees' compensation for training time were included (Eurich, 1985).

Less measurable than formal training, on-the-job training has been estimated to cost $180 billion annually (Carnevale, 1986). The amount spent for formal training and on-the-job training may total the huge sum of $210 billion per year. To understand how the amount spent for training compares with expenditures in other kinds of industries, consider the following estimates: The aerospace industry costs $80 billion annually, formal education costs $238 billion annually, and the health industries cost $355 billion annually (ASTD, 1986; Carnevale, 1986). Whatever the exact amount, it is clear that employers are running a massive program that is a
critical addition to an employee's elementary, secondary, and postsecondary education.

The literature on human capital theory and the scope of employer training helps bring into focus the following:

1. The postindustrial economy requires an increasingly skilled labor force (Ginzberg & Vojta, 1981).

2. Efforts to maintain and increase the skill of the workforce are falling squarely on education and employer training (Eurich, 1985).

3. The substantial investment in human capital through employer training makes it imperative that the quality of that training be maintained (Lynton, 1984).

To insure quality and a return on the investment in training, organizational departments responsible for learning experiences must approach training activities in a deliberate way. One such way was described in a program-development model formulated by Brinkerhoff (1987). Brinkerhoff's model was developed around key information-gathering points from which informed decisions can be made about the merit and worth of learning-program activities. Brinkerhoff described these points in six stages. Stage I asks: Is there good reason to do some training? Could it make a difference? Stage II asks: Is the training design good enough to implement? Stage III asks: Is the design installed and working? What problems are occurring? Stage IV asks: What reactions in terms of skills, knowledge, and attitude (SKA) changes have occurred? Stage V
asks: Have reaction (SKA) changes lasted? Is expected use occurring? Stage VI asks: What difference did it make? Was it worth it?

Evidence of the merit of training can be gathered from Stages II, III, and IV. Meritorious training will have a good training design, will be effectively conducted, and will result in learning. Stages I, V, and VI are concerned with the worth of training. Can the learning be used in the job environment? Do the resultant changes in skills, knowledge, and attitudes make a valuable contribution to the organization? If there are no evaluation activities at Stages I, V, and VI, information with which to assess worth is not available.

Although evaluation is usually thought of as a way to assess program operations and effects retrospectively, the most important phases of the cycle described by Brinkerhoff and of similar cycles, such as those described by Kirkpatrick (1978) and Minick and Medlin (1983), occur at the beginning. It is during the planning and investigative stages that evaluation activities may have the greatest effect. If, for example, the corporation does not need the knowledge and skill the training program is planning to deliver, it will be impossible to produce a program that has worth. Similarly, if the problem to be solved is an environmental one rather than one involving skills, attitudes, or knowledge, no amount of training will produce the desired results. Plans to gather information at various stages during the cycle should be made when the training
program is designed. Information gathering or evaluation should be identified with that program from beginning to end (Bass & Vaughan, 1966).

Systematic evaluation at many stages in the training process helps trainers make decisions about training needs, training plans, training operations, and training results. Information about what is needed, what is working, and what has happened as a result is basic to the design and delivery of programs that have both merit and worth.

Investments in training, education, and development, like investments in other resources, are expected to produce a return of value to the corporation or organization. To demonstrate that value, training practitioners must evaluate both the merit and the worth of their programs. Information about merit provides evidence that something has been done well; information about worth provides evidence that something has had valuable results. There is no value in doing something well that was not worth doing at all. "Good HRD [human resource development] must seek both merit and worth. Merit is necessary, but not sufficient for worth" (Brinkerhoff, 1987, p. 34).

The discipline of evaluation research has evolved rapidly in the past 20 years, primarily in response to the need for evaluating government-sponsored social and educational programs. Evaluation literature has offered a variety of evaluation models that can be applied to the corporate setting. However, despite the relationship
between evaluation activity and the design and delivery of learning programs, and despite the proliferation of literature and models, there is widespread agreement that the use of evaluation lags far behind the encouragement it has been given in professional literature (Brandenburg, 1982; Carlisle, 1984; Dopyera & Pitone, 1983; Swierczek & Carmichael, 1985).

This lag is most noticeable where evaluation efforts concern complex evaluation activities—those that assess the worth of the learning experience. Complex evaluation activities ask the questions: Would training make a difference? Do participants use newly acquired skills, knowledge, and attitudes (SKA)? Have SKA changes lasted? What difference do the new SKAs make? What contributions do the new SKAs make to corporate goals?

A number of recent evaluation studies have shown evidence that complex evaluation activities to assess worth are inadequate. The Bureau of National Affairs (1985) surveyed 159 organizations throughout the United States and found that 7 out of 10 trainers conducted some kind of training evaluation. Almost all (96%-99%) used participation reaction, slightly more than half (64%-69%) measured changes in job performance, and few (25%-27%) evaluated the effects of training on company goals. In a similar survey conducted by Louis Allen Associates, it was found that 9 out of every 10 trainers used participation reactions, whereas only 1 out of 4 evaluated the transfer of training to the job and less than 7 in 100 assessed the effect of training on company profits (Miller, 1984).
Brandenburg (1982) reported that the most frequently used evaluative techniques were the "smile" indices, i.e., participant reactions and comments. Cognitive and performance-based outcomes were used less often, whereas the least used techniques were those that measured the lasting effects of training back on the job.

Most typically, then, training programs are evaluated through the "happiness" ratings of the participants. Even by optimistic estimates, less than 50% of all companies evaluate the transfer of training to the job, and less than 20% use any productivity measures to judge the worth of training. Both the transfer of training to the job and productivity measures are complex evaluation activities. As evaluation activities become more complex, fewer training departments employ them. Rather than using the effect of training to judge the worth of training programs, training departments are most often judged through activity reports. Such reports chronicle how many workshops have been conducted, how many participants there were, and how many learner-contact hours were generated. Reports of this type often carry the message that maintaining a level of activity is important, when in fact only results are important (King & Roth, 1981; Larson, 1985).

Corporate learning, at present a $30 billion industry, is large and likely to grow larger. Economic and technical changes will, in effect, fuel the necessity for training and development in the workplace (Carnevale, 1986). Despite the size of the investment in corporate learning and despite its almost certain future demands,
careful planning and complex evaluation activities often are conspicuous by their absence. The investments in and potential of organizational training evaluation are too great to allow a "head in the sand" approach. Competent planning and evaluation are the foundations needed to meet organizational goals.

Purpose of the Study

The rigorous approach to proving the worth of training and the systematic evaluation of training programs have received enthusiastic support at professional conferences and in professional journals. Just glancing through any training publication, one finds a number of articles that argue persuasively for training-program evaluation. Using "shoulds" and "oughts," the general message is: Every training program should be evaluated and every trainer ought to be engaged in a series of evaluation activities (Brethower & Rummel, 1979; Bunker & Cohen, 1978; Kirkpatrick, 1977; Larson, 1985).

In spite of the proliferation of evaluation literature and the growing investment being made in training activities, it appears that the evaluation component remains a totally underdeveloped element. Without attention being given to evaluation activities that assess both the merit and worth aspects of training, the design of effective and efficient programs is initially flawed. In a growing training and development industry, this casts suspicion on
whether organizations are getting the return on investment that they could be realizing from the training activities they sponsor.

This study addresses possible reasons why professionals responsible for training activities are not involved in systematically evaluating both the merit and worth of those programs. The results of this study will enable practitioners to gain insight into the reasons for the gap between professional literature and actual practice regarding training-program evaluation.

Organization of the Study

This dissertation comprises five chapters. The background for the study as well as the importance and purpose of the study were outlined in Chapter I.

Chapter II contains a review of literature related to the topic under investigation. Evaluation models and current practices in training evaluation are examined. The need, value, and feasibility of gathering complete and complex evaluation data are discussed. The purpose of the literature review is to highlight theories, trends, and practices that provide the theoretical framework for understanding the role of training evaluation and its relationship to the need for, value of, and feasibility of gathering complex evaluation data.

The design of the study is presented in Chapter III. Included are descriptions of the development of the research instrument, the
sampling procedure, the collection of data, and methods used to analyze the data.

The results of the data analysis are presented in Chapter IV. Statistical procedures are described and the results discussed.

Chapter V contains a summary of the study, conclusions, and recommendations for further study, and implications of the research findings.
CHAPTER II

REVIEW OF RELATED LITERATURE

The substantial investment in human capital through employer training, as discussed in Chapter I, provides the background for the review of related literature. In this chapter, significant evaluation models that measure both the merit and worth of training are highlighted, and current practices in training evaluation are discussed. Next, the need for and value of evaluation and the feasibility of gathering evaluation data in the organizational setting are considered. Finally, the chapter ends with the research questions, a statement of the hypotheses tested in the study, the assumptions, and limitations of the study.

Training-Evaluation Models

The value of human capital and the scope of employer training make it imperative that the appropriateness of any kind or amount of training not be accepted uncritically (Lynton, 1984). Barton (1982) supported this contention by asserting that training should not be promoted on the premise that it can perform miracles. Employers need facts to support decisions concerning how much and what kinds of training are most effective.
The efforts of training evaluation are directed toward appraising the quality of existing training and providing a sound basis for decision making. Evaluation is a means toward the end of improved quality in training.

Hamblin (1974) posed three questions he thought were central to the improvement of training-evaluation practice: How can the criteria by which training is evaluated be improved? How can the methods for gathering information on the basis of which training is evaluated be improved? How can evaluation be used to improve training? These three questions are interrelated and deal with Barton's (1982) and Lynton's (1984) concerns about the basis on which to make decisions about the quality of training.

Several training-evaluation theorists have proposed systematic models for identifying areas or data levels within the training process to which both trainers and management might look for evaluation and decision-making information (Brinkerhoff, 1987; Hamblin, 1974; Hesseling, 1966; Kirkpatrick, 1977; Warr, Bird, & Rackham, 1970). The benefits of these models are that they provide a means for improving training and a way to examine the transfer of training into the organization for which the training was designed.

The most widely accepted approach to training evaluation is attributed to Kirkpatrick. In 1959 and 1960, he outlined his approach to evaluation in four serial articles in the Training Director's Journal. He has presented this approach in several forms since then (1977, 1978). Kirkpatrick suggested four areas to
explore in evaluating training; a trainer may choose to examine one area or all four. These areas are discussed in the following paragraphs.

1. Reactions. Reaction data include how the participant feels about the training. Such data are, in a sense, a measure of "customer satisfaction" (Kirkpatrick, 1978). Other writers have referred to this level of data as a "happiness rating." Data related to this area would probably be gathered at the end of the training day or program. Reaction sheets may ask participants to rate format, instructor presentation, use of media, and usefulness of content, providing the easiest and least costly data to collect. Hamblin (1974) pointed out that, although they are often maligned, happiness ratings are important because the trainees' positive reactions are central to the openness of their learning attitudes.

2. Learning. A second data level, learning, includes the areas of knowledge, skills, and attitudes. Data collected in this area include a demonstrated change in amount of knowledge, ability to perform a new skill, or behavioral responses that indicates a change in attitude. Any change in knowledge, skill, or attitude indicates that learning has taken place.

3. Behavior/performance. On-the-job behavior dealing with the transfer of learning is the third data level for evaluation. Data on change in behavior involve performance appraisals--ideally, on a before-and-after basis. These data may be collected from
self-reports, interviews with trainees, supervisor evaluations, or peer appraisals. Performance or behavior data should be gathered some time after training to demonstrate retention.

4. Results/impact. This data level is often referred to as the "bottom line." If trainees learned a skill and transfer it to the job, what difference does it make to the organization? In an industrial setting, the results might be a measure of employee turnover or an improvement in sales. In a correction setting, the results might be fewer rule violations as indicated by a decrease in "tickets" issued. Kirkpatrick (1978) and Hamblin (1974) both emphasized the difficulty of learning about the ultimate effect of training. The problem is often one of separating variables as organizational changes are traced back to training events. Hamblin believed that in many cases this separation is so difficult that it is not worth the attempt.

Brethower and Rummler (1979) added strategies for gathering data at four levels of evaluation. In addition, they offered a summary of several research designs (control group, reversal, multiple base line, and "before and after" measurement) and their applicability to the "real" world.

Hesseling (1966) presented quite a different classification system for the evaluation of training. He called it a typology of evaluation and combined two points of reference: "for whom" the evaluation is created and "by whom" the evaluation is conducted. The typology provides a matrix, with the horizontal axis being
assessment for whom and the vertical axis being assessment by whom. In this way Hesseling arrived at 25 types of evaluations in the cross-tabulation. The typology is intended to bring together in one conceptual framework the pragmatic, everyday activities of judgment, appraisal, and reporting with the more sophisticated evaluation research activities. To make the typology operational, Hesseling added a third dimension, that of planning, action, and fact finding, all of which can be applied in each cell of the matrix. In this way evaluation is used in its broadest sense to influence the training activities of planning, action, and fact finding.

Warr et al. (1970) developed another model offering a four-stage evaluation process, which is commonly known as the CIRO (context, input, reactions, and output) classification. The context evaluation stage is the one during which the question, What needs to be changed? is considered. In this stage immediate, intermediate, and ultimate objectives are formed. During the input evaluation stage, the trainer deals with the question, What procedures are most likely to bring about change? In this stage the resources most likely to accomplish the objectives are evaluated. Reaction evaluation is defined as activities used to gain information about trainees' expressed reactions. This type of evaluation is distinguished by its totally subjective nature. Output evaluation, by contrast, is the objective measure of training results. This stage may be applied at any time during the process, depending on
when an appropriate measurement must be taken. The stages in the CIRO framework are not distinct and sequential but are interactive and emphasize the trainers' need for information to help decide how resources can be used to increase the effectiveness of training and thus help attain organizational objectives.

Anthony Hamblin (1974), a British management expert, developed a more complicated taxonomy of data levels. He suggested looking at both the causes and effects of training, reactions, learning, changes in the organization, and changes in the achievement of ultimate goals. Hamblin suggested that these five levels depend on one another, training leading to reactions, leading to learning, leading to changes in the organization and achievement of ultimate goals. In this chain of effects, the evaluator's task is to "discover whether the chain held through its links" (p. 15).

Bakken and Bernstein (1982) used the work of Kirkpatrick (1978) and Hamblin (1974) to develop a systematic approach to designing evaluation of training. Based on Hamblin's five general objectives of training (reaction objectives, learning objectives, job-behavior objectives, organizational objectives, and ultimate objectives) and Kirkpatrick's four types of outcomes (reactions, learning, job performance, and organizational impact), training can be planned and measurements selected to fill the needs of various decision makers. Bakken and Bernstein included the important step of identifying the decision makers who might need or later seek information about the effectiveness of a particular training program. Once the
decision makers have been identified, the evaluation plan should provide the desired information to the appropriate decision maker.

A six-stage evaluation model formulated by Brinkerhoff (1987) and described in Chapter I of this dissertation combines a program-improvement focus with a results-oriented or "bottom-line" approach and an emphasis on utility. Brinkerhoff believed that the evaluation approaches described by Hamblin (1972), Kirkpatrick (1978), and Warr et al. (1970) focused on results and largely ignored the program-development and program-improvement needs of trainers. Intended for use by all training practitioners, regardless of their level of sophistication, the six-stage model uses problem-solving aspects of educational evaluation and employs, but does not rely solely on, verifiable scientific measurement. It is not a theoretical or technical approach but rather a practical, conceptual model. Stage I asks: Is there a good reason to do some training? Stage II asks: Is the training design good enough to implement? Stage III asks: Is the design installed and working? Stage IV asks: What reactions in terms of skills, knowledge, and attitude changes have occurred? Stage V asks: Have changes lasted? Stage VI asks: What difference did the training make: Was it worth it?

Brinkerhoff's (1987) six-stage model provides two key benefits. One is the distinction between merit and worth. Identified in the model are activities used to evaluate merit (Stages II, III, and IV).
and those used to evaluate worth (Stages I, V, and VI). Merit tells how well something was done; worth describes whether the "doing" had any valuable results. Being clear about that distinction may help practitioners avoid doing well what was not worth doing in the first place. A second benefit of Brinkerhoff's model is that it provides for the evaluation of both training outcomes and training processes.

Typically, these models and variations of them are based on similar processes: needs assessment, program development and implementation, program monitoring, and measurement of results. Each, to varying degrees, represents a total system incorporating important phases that should be part of the well-planned training program. These types of models have several advantages, some of which are as follows:

1. They provide a framework for planning and evaluation, emphasizing that evaluation is an important part of the total system, not just a summative activity.

2. They direct attention to the "fit" that ought to exist between corporate goals and training objectives.

3. Most are a closed-loop system that uses feedback to modify the program continuously. Training programs are viewed as unfinished products to be continually modified (Woodington, 1980).

Current Practices in Training Evaluation

Very few training professionals either have no opinion regarding evaluation or do not support the evaluative function.
Although there are many models for evaluating training, none is universally accepted (Brandenburg, 1982). In 1961, McGehee and Thayer placed the evaluation of training in the same category as Mark Twain did the weather: Everybody talks about it, but nobody does anything about it. Large amounts of time, effort, and money are spent on developing various types of training programs, which indicates that the need for such training is recognized. A review of the literature suggested, however, that the evaluation of such training has been given less than adequate attention.

In 1953, French (cited in Woodington, 1980) found that only 1 company in 40 had any systematic approach to the evaluation of supervisory training. Findings of a 1961 study of evaluation practices in management training showed that most training departments spent 1% to 5% of their time on evaluation activities, one fifth spent less than 1%, and another one fifth spent 6% to 10% of their time on evaluation (Shafer, 1961).

Kirkpatrick (cited in Catalanello & Kirkpatrick, 1968) surveyed the extent of training evaluation by sending questionnaires to 154 firms that had indicated an interest in evaluation. He found that 78% of the respondents used participant reaction as an evaluation measure, 50% measured skill learning, 54% evaluated performance, and 45% evaluated results. Catalanello and Kirkpatrick observed that those evaluations which looked at behavior and results were
subjective and superficial, and that evaluation of training was still in its infancy.

Sullivan (1970) surveyed the management-training evaluation procedures of 50 companies randomly selected from the Fortune 500 list and representing a range of product industries. He found that (a) there was a substantial gap between evaluation theory and practice, (b) the relationship between training and behavior was difficult to measure because of variables other than training, (c) evaluation of management training tended to be superficial and subjective, (d) the primary criterion was behavior change, and (e) the primary reason for poor evaluation practices was lack of know-how. Sullivan also found that the primary approach was reaction evaluation, followed by pre-/posttest changes in learning.

In 1978, Clegg replicated Sullivan’s study, enlarging the questionnaire and sending it to the original 50 companies. Clegg found that in ranking evaluation indicators, the top three were: change in performance on the job, students’ reaction to training, and changes in knowledge possessed by students. Other evaluation activities reported included surveying management’s opinion of training results, changes in organizational impact, and documenting the cost of training. Clegg’s recommendations included better training for evaluators in the methods and theory of evaluation.

In a 1980 article, Brown discussed an unpublished study by Smeltzer, who examined the evaluation practices in 250 companies. Smeltzer found that only 12% of the companies evaluated management
and supervisory training at all, and evaluations that were done were of the reaction type.

Brandenburg (1982) also compared evaluation activities across organizations. He collected data from two professional training groups on the functions of evaluation, data-collection techniques, and the skills needed to conduct successful evaluations. One purpose of the study was to determine how frequently 12 data-collection techniques were used when the training programs were being evaluated. The most highly ranked techniques were objective questionnaires or surveys, multiple choice or similar achievement measures, task performance measures, and observation or anecdotal records. Both cost analysis and essay or open-ended achievement exams were ranked very low. Later on-the-job performance measures and indirect follow-up studies were found to be done once in a while. Brandenburg concluded that techniques used most often were performed while participants were on-site in training and that these techniques were efficient, short, and not labor intensive. More complex techniques, requiring more time or personnel, consistently received lower priority and were used less often. Thus, most evaluation efforts were focused on the immediate aspects of training and did not reflect the context of the total organization.

A study published in 1985 by the Bureau of National Affairs shed considerable light on current evaluation practices. The Bureau surveyed 159 organizations through the United States concerning all
aspects of their training and development programs, including their evaluation practices. Of the 159 organizations represented by the respondents, 48% were manufacturing firms, 23% were nonmanufacturing businesses, and 29% were nonbusiness establishments such as health care organizations, educational institutions, and government agencies. In terms of size of workforce, 50% of the participating organizations were classified as small (fewer than 1,000 employees) and 50% were considered large (workforce numbering 1,000 or more).

Report findings showed that for nonmanagement training, 7 out of 10 trainers conducted some kind of training evaluation. Of those who conducted such evaluations, 96% relied on participant reaction, 69% monitored training success by measuring change in the employee's job performance, and 25% evaluated training programs according to changes in measures of the company's productivity (profit, sales, employee absences, turnover, and so on). Results showed some differences in whether or not evaluations were conducted, depending on the type and size of the organization. Sixty-three percent of the manufacturing organizations evaluated nonmanagement training, 81% of the nonmanufacturing business organizations evaluated nonmanagement training, and 67% of the nonbusiness organizations evaluated nonmanagement training. By workforce size, 75% of the large companies evaluated their nonmanagement training programs, as compared with 63% of the smaller firms.

Statistics for the evaluation of supervisory and management training were very similar. Seventy-one percent of the surveyed
companies providing supervisory and management training conducted formal evaluations of these training programs. Nearly all (99%) of the companies doing evaluations relied on participant reactions, 64% looked for improvement in job performance, and 27% evaluated the effect of training on measures of the company's productivity.

Nonmanufacturing firms (88%) were much more likely than nonbusiness (68%) and manufacturing (64%) concerns to evaluate their management programs in some way. About three fourths (76%) of the large organizations conducted formal management-training evaluations, as compared with slightly less than two thirds (64%) of the smaller firms (Bureau of National Affairs, 1985).

Current literature indicated that most training evaluations measure the degree to which trainees and their supervisors accept the program. Acceptability alone, however, does not prove the effectiveness of training efforts. Of evaluations that go beyond trainee or supervisor reactions, the majority evaluate trainee performance against predetermined criteria immediately after training.

Numerous writers offered testimonials on the importance of training evaluation and gave examples of methods for gathering data on the worth of training programs. Published evaluations have a role in demonstrating what has been done and how complete or incomplete the science of evaluation is. The few examples offered here are representative of the whole in that they all advocated a
complex level of evaluation activity (i.e., transfer of training and productivity measures).

Urban, Ferris, Crowe, and Hiller (1985) reported on a program-evaluation effort of a large-scale management training and development program in a major oil company. Their evaluation approach employed several different methods and provided for the examination of many criteria from which conclusions were drawn as to the program's effectiveness. Positive participant reaction, postprogram promotion, pay-grade increases, reduced turnover rates, and cost-benefit measures were all used to support the effectiveness of the management training and development program. Urban et al. concluded that, in the current uncertain economic climate, fewer resources will be allocated to those programs whose worth cannot be justified.

Swierczek and Carmichael (1985) used the Planning and Organization for Effective Supervision workshops, offered to public agencies in Florida by the University of South Florida, to illustrate the quantitative and qualitative aspects of evaluating training. Their approach combined qualitative participant responses to open-ended questions with the quantitative measurement of training effectiveness gained through pretest and posttest data. The writers concluded that using both approaches would help trainers improve training programs and enable them to assess learning and provide management with a measure of the transfer of learning to the workplace.
Advocating a "bottom-line" approach to the evaluation of training, Keiley, Orgel, and Baer (1984) suggested using direct performance measures that reflect the actual effect of training on productivity. They advocated using productivity figures that companies already collect, so that training, which is supposed to improve that productivity, can be evaluated. Giving several examples, they showed how on-the-job performance measures were used to assess the costs and benefits of training.

In an article designed to convince training practitioners that they cannot afford to ignore the issues of training effectiveness and worth, Bunker and Cohen (1978) refocused attention on the basic methodological issues of evaluation. They offered some explanations for existing inadequacies in training evaluations and demonstrated how these inadequacies can be overcome. Bunker and Cohen contended that, although a half-hearted evaluation effort may be better than none at all, adding some rigor and control to evaluation efforts is well worth the effort.

Some attention has been given to using assessment centers to evaluate training. In a typical assessment center, training is evaluated by having trainees work through simulations that mirror work situations. The simulations are constructed so that the results indicate whether or not trainees are able to apply what they have learned in training. As an evaluation method, assessment centers do not provide a complete indicator of training
effectiveness because they cannot produce evidence of the trainees' actual use of the skill on the job. However, this method provides an excellent indication of acquisition of the new skill and in that sense is a step beyond many evaluations that deal only with acquisition of new knowledge (Byham, 1982).

Consistently, writers giving examples of training evaluation in specific situations emphasized the importance of measuring the worth of training. Without such a measure, there is little or no evidence to show a return on investment. Seen as a liability or expense rather than an asset or a revenue-generating investment, training programs may have little credibility in a profit-oriented environment (Carnevale, 1986; King & Roth, 1981).

Need for, Value of, and Feasibility of Training Evaluation

Whether or not and to what degree a training department develops training-evaluation plans hinges on the issues of need, value, and feasibility. The first two, need and value, are related to the question, Why evaluate? It must be clear that there is a need to evaluate. That need is associated with the discrepancy that exists between conditions of nonevaluation and conditions of evaluation. For the need to be met, the training manager would have to choose the condition of evaluation. The second issue is value. There must be some payoff with regard to the evaluation outcomes. The third issue, feasibility, is related to the question, Why not
evaluate? It deals with deterrest and how possible it is to achieve the conditions that allow for program evaluation.

Need for Training Evaluation

Both Brandenburg (1982) and Brinkerhoff (1987) asserted that evaluation to improve training programs is a primary need. Brinkerhoff, in particular, made the point that training cannot be done well without accurate information. Information about the context in which training is conducted, as well as the evaluation of needs, plans, operations, and effects, is necessary to develop and refine training. Collecting data from two professional training groups, Brandenburg found that the top four functions of evaluation, as these groups saw them, were to improve the training programs, to provide feedback to program planners or management, to gain knowledge of employees' skill levels, and to provide feedback to program participants.

Many other writers have addressed additional aspects of the need for program improvement. McGehee and Thayer (1961) specifically pointed out that evaluation allows the trainer to compare various means or techniques of training to determine if any one or a combination is best to achieve desired results. Newstrom (1978) suggested that evaluation allows the trainer to assess the achievement of training objectives and the effectiveness of the trainer. In addition, evaluation can provide information to
make a decision about whether a particular program should be repeated (Smith, 1980).

Another need, other than to improve programs, is to prove their effect and worth (Brinkerhoff, 1987), that is, to determine what kind of contribution the program has made to the organization. Zenger and Hargis (1982) pointed out that, because of large expenditures on training, management wants to see evidence of a return on the investment. Quinn and Karp (1986) added that training results are measured to justify conducting the training and to provide evidence that the training contributed to the company's goals and objectives--or, in McGehee and Thayer's (1961) words, "Are the dollars being spent on training producing the results needed by the organization?" (p. 257). Kelley et al. (1984) suggested that developing evaluation methods to identify accurately and thoroughly the benefits and costs of training may become the most important contribution of training evaluation in the 1980s.

In a note of caution about claiming "bottom-line" results, Zemke (1982) wrote: "The contention that training and [human resource development] efforts can prove their bottom-line results may be the greatest myth ever concocted by the profession" (p. 26). Maintaining the posture that training is part of the total action plan, Zemke reminded trainers that, like every other staff function, training helps line managers and line workers attain superior performance. Taking sole credit for an increase in production, profits, and so on, can contribute to unproductive conflict.
Training directors need to be realistic about what is and is not possible through training.

Value of Training Evaluation

Throughout the literature, the ultimate payoff to training programs that include complex evaluation activities was seen in terms of credibility and survival. In a 1979 ASID survey, respondents were asked, "What is the most important behavioral requirement for success as a training and development professional?" More than 40% of the 12,089 participating trainers identified credibility as the most important requirement. The second most important requirement was flexibility, identified by 18% as being important (Clement, Walker, & Pinto, 1979). Clearly, one of the critical factors in the organizational success and support of training is the ability to maintain and build credibility.

Credibility in the organizational environment exists when others view with confidence the trainer's ability to get the job done. Organization managers trust that credible trainers can contribute to the organization's goals because they have demonstrated the ability to do so. Some see credibility among middle managers as an important avenue to getting support from upper management (Clement et al., 1979).

According to King and Roth (1981), the key factor in gaining credibility is to have a direct influence on organizational goals.
and, through evaluation, be able to demonstrate that effect. When trainers do not provide good evaluation research, management support is based primarily on faith and good will, both of which can change with the economy. Experience has shown that, in times of economic downturn, training budgets are among the first to be cut (Chabotar, 1977; Kelley et al., 1984; Zenger & Hargis, 1982).

Another value of credibility is survival—not just if the training program survives but how well it survives. Not being viewed as credible in terms of playing a contributing role in goal achievement can have two major consequences for the training manager. First, the training director may not be given sufficient resources in terms of staff and budget; second, he/she may play a minor or secondary role in the corporate organization (King & Roth, 1981).

As manager of a department, the training director is competing for corporate resources. If the training director cannot demonstrate the return on investment that other corporate managers do, he/she may be unable to compete effectively for financial support (Larsen, 1985). In most corporations, power is associated with budget size and control over allocating the budget. The more power one has, the more money he/she will have to spend and the more authority in deciding how it is spent (Zemke, 1984a).

A readership survey conducted in 1984 by Training magazine revealed that "trainers who reported that their budgets increased over the previous year, also reported that their top management sees
training as important and the greater the budget increase the higher the rating of training's importance in management's eyes" (Zemke, 1984a, pp. 73-74). If the corporation is willing to invest in training and development programs, it will commit itself financially, making a statement about the value of training to the entire corporation (Eurich, 1985).

The role the training director plays in the operation of the organization may also indicate the value training has to the organization. The role the training director develops may be a consequence of the perceived effect of training on corporate goals. Typically, training directors are not part of the corporation's decision-making process. They generally do not sit on policy-making committees, are not part of the management team operationally, and therefore are not part of strategic planning organizationally (King & Reth, 1981).

Some writers believe an indicator of the extent and success of a training program is the position of the person managing that program. The higher the training director's position in the organization, the more this reflects the commitment to training. Appointing a corporate-level executive to head the training function makes a statement to all managers about the importance of the training function in the organization (Eurich, 1985).

A 1985 readership survey conducted by Training magazine led Zemke (1985) to draw some conclusions about characteristics of
training programs that are flourishing in the organizational environment. The survey was an attempt to determine how training practitioners approached the design, development, delivery, and evaluation of training programs. The participants responded to 23 statements, 20 of which reflected specific practices associated with a systematic approach to training, much like the ones suggested by Brinkerhoff (1987). Zemke reported that there seemed to be a significant relationship between certain secondary indicators (budgets, respect, and management participation) and a general systematic approach to the design, development, delivery, and evaluation of training. He concluded: "Those who practice a systematic approach fare better in their organizations than those who don't" (p. 108).

Feasibility of Training Evaluation

The issue of feasibility is related to the question, Why not evaluate? In spite of the abundance of material concerning proper evaluation design, suggested techniques remain largely unused (Blakeslee, 1982). The literature contained many reasons for not evaluating; some were the opinions of a single writer, whereas others had been collected through research.

Several research studies on management-training evaluation offered insight into why effective evaluation techniques are not used (Clegg, 1978; Shafer, 1961; Sullivan, 1970). In 1961, Shafer conducted a study to determine both the theoretical and actual
practices in evaluating formal management training in industry. His research design consisted of a questionnaire sent to 158 large companies. One of Shafer's findings was that the major deterrent to effective evaluation was that evaluation techniques were too difficult to apply in production situations and were too subject to variables that could not be controlled. Blakeslee (1982) echoed many of the same thoughts; he pointed out that the daily demands of the work environment make it almost impossible to develop valid and reliable evaluation designs.

Sullivan (1970) completed a field survey of 50 companies to determine what management-evaluation practices were being used in American industry. His conclusions regarding the difficulties involved in evaluation supported Shafer's findings. However, Sullivan added that the primary reason for poor management-training evaluation practices was the lack of evaluation know-how. Woodington (1980) supported this finding: "The chief problem is not one of willingness to accept evaluation--but one of being able to do it well" (p. 328).

In 1978, Clegg replicated Sullivan's study. Clegg reported four reasons for not evaluating management-training programs:

1. The individual responsible for evaluation does not have enough time.

2. The responsible officials do not know what to evaluate because of unclear objectives.
3. Some responsible officials do not know how to go about evaluating.

4. Since some companies do not require evaluations, no attempt is made to evaluate.

Minick and Medlin (1983) supported the contention that the time and effort needed to implement effective evaluation often are usurped by other production pressures. They stated that the basic commitment is the time involved for information gathering and processing.

Clegg's finding that perhaps no evaluation was made because none was required received some support in the literature. Deming (1982) believed that evaluation necessitates having an evaluator with the ability to complete an assessment and a decision maker with the conviction that it is important to know whether training has had the desired effect on human performance.

Concluding a discussion on the evaluation of training, McGehee and Thayer (1961) expressed confidence that training specialists will spend more time in training evaluation if "management accepts training as a management tool and asks loudly what training has contributed to the achievement of organizational goals" (p. 285).

In summary, Wolfe (1973) listed several deterrents to evaluation: It costs too much, intervening variables cannot be controlled, measuring devices are not available, the staff is not qualified, the statistical work is cumbersome and complicated, it is doubtful that a relationship between training and results can be
established anyway, there is not enough time or personnel, and the evaluation results would be too theoretical. Recognizing all of these as dilemmas, Brandenburg (1982) submitted that if there were no problems in evaluation, someone would be assigned to do it. He suggested that "if follow-up information could be demonstrated to have important organizational consequences, then training managers would find means to accomplish it" (p. 18).

The purpose of the present study was to investigate the discrepancy between the emphasis on training-program evaluation in the literature and actual practice in the field. In investigating this discrepancy, there is value in understanding that behavior is often a function of its consequences. Simply stated, people are more likely to initiate or repeat an action if its consequences are desirable, just as they are less likely to initiate or repeat an action if its consequences are negative or unpleasant (Rosenbaum, 1982). In an organizational setting, managers are more likely to perform an activity, the more desirable they perceive the reward of that activity to be (Hampton, 1972).

Managers in a corporate environment are evaluated in terms of how effectively and efficiently they can manage their resources to have the greatest positive influence on identified goals. Good managers are able to manage resources for maximum effect better than others. Good managers are rewarded for their positive effect on corporate goals because they communicate that contribution to key
decision makers in the corporation (King & Roth, 1981). It is reasonable to conclude that training directors are also more likely to perform an activity if they perceive the reward of that activity to be desirable.

In very practical terms, perhaps complex planning and evaluation activities are lagging because they are not required or rewarded by upper management. Is it because the more complex evaluation activities are given relatively low priority by upper management that evaluation results often reflect only the most immediate aspects of training (participant reactions)? If training directors are not required to show or are not rewarded for demonstrating the worth of training, why should they spend time and resources on complex and complete evaluation activities? Could the discrepancy be due in part to a difference between the perceptions of upper-level managers and training directors concerning the value of, need for, and feasibility of gathering evaluation data on training, education, and development activities?

The writer's intention was to examine the perceptions of two of the stakeholders in the training process: upper-level managers (decision makers) and training directors. Specifically, their perceptions were compared as to (a) the value each group placed on complex training-evaluation activities, (b) what each saw as the need for such activities, and (c) what each understood to be the feasibility of gathering complex evaluation data.
Research Questions

This study was designed to answer the following research questions:

1. Is there a discrepancy between upper-level managers' and training directors' perceptions regarding the degree to which the evaluation of training, education, and development activities is needed in the organization?

2. Is there a discrepancy between upper-level managers and training directors regarding the value they place on the evaluation of training, education, and development activities?

3. Is there a discrepancy between upper-level managers' and training directors' perceptions of the feasibility of gathering evaluation data on training, education, and development activities?

4. Is there a relationship between the discrepancies in upper-level managers' and training directors' perceptions of the value of, need for, and feasibility of gathering evaluation data and the degree to which evaluation activities are accomplished?

Hypotheses Tested

The following directional hypotheses were formulated to investigate the research questions posed in this study:

Hypothesis 1: The frequency with which the training department conducts training-evaluation activities is related to the extent of positive congruence in the perceptions of the upper-level managers and training directors concerning the value these evaluation activities would have in generating management support.
Hypothesis 2: The frequency with which the training department conducts training-evaluation activities is related to the extent of positive congruence in the perceptions of the upper-level managers and the training directors concerning the need for these evaluation activities to help produce more effective training.

Hypothesis 3: The frequency with which the training department conducts training-evaluation activities is related to the extent of positive congruence in the perceptions of the upper-level managers and the training directors concerning the feasibility of accomplishing those activities in their particular organizations.

Hypothesis 4: Where there is disagreement in perceptions between the upper-level manager and the training director concerning the value of the evaluation activity to generate management support, it is the positive perception of the upper-level manager that will be related to the frequency with which the activity is carried out.

Hypothesis 5: Where there is disagreement in perceptions between the upper-level manager and the training director concerning the need for the evaluation activity to produce more effective training, it is the positive perception of the upper-level manager that will be related to the frequency with which the activity is carried out.

Hypothesis 6: Where there is disagreement in perceptions between the upper-level manager and the training director concerning the feasibility of accomplishing the evaluation activity in their particular organization, it is the positive perception of the upper-level manager that will be related to the frequency with which the activity is carried out.

Assumptions

The writer made the following basic assumptions in conducting this study:

1. That the respondents and researcher had a mutual understanding of the following terms used in the study:
Training: An intervention taking a variety of forms including but not limited to classroom lectures, workshops, self-directed learning, etc., whose purpose is to help employees acquire or improve on a skill, knowledge, or attitude.

In-house training: Organized training developed and/or conducted by the sponsoring organization's training department for the exclusive use of the organization's employees.

On-the-job training: A training intervention that occurs on the work site and is generally conducted by a supervisor, lead worker, or fellow employee.

Training program evaluation: The systematic inquiry into training contexts, needs, plans, operations, and/or effects to determine what has happened as a result of training (Brinkerhoff, 1981).

Ad hoc committee: A group of people brought together for an express purpose, such as to conduct evaluation of training programs.

In-house evaluation experts: Persons employed full time by the organization, who have as a part or all of their job the task of conducting evaluations.

Outside consultant: A person hired on a job-by-job basis by the organization for the express purpose of delivering a training program, conducting a needs assessment, or conducting a program evaluation.
2. That the respondents and researcher had a mutual understanding of the following concepts used in the study:

**An activity that would help generate management support:**
This was used by the researcher as the value of an evaluation activity. The respondents were asked in the written questionnaire to judge if each of the 11 evaluation activities would produce information that would help generate management support, usually evidenced by increased budget, staff, and/or esteem in the organization.

**An activity that would help produce more effective training:**
This was used by the writer as the need for an evaluation activity. The respondents were asked in the written questionnaire to judge if each of the 11 evaluation activities would produce information to help make the training activity more effective in terms of the learner acquiring new or improved skills, knowledge, and attitudes.

**An activity that would be feasible in their organization:**
This was used by the writer to determine the respondents' perceptions of whether or not conducting any of the 11 evaluation activities was possible in their circumstances.

3. That the respondents and researcher had a mutual understanding of the following methodologies being investigated:

**End-of-training evaluation concerning trainee satisfaction with the training experience:**
Generally called "reactions or
smile indices," these evaluations by trainees usually respond to the content covered, presentation techniques, probability of the trainees using the new knowledge/skills back on the job, and any other questions considered important by the trainer or comments considered important by the trainee.

Specifically formulating training outcomes to reflect organizational need: These outcomes are generally stated in objective form, which in this case would reflect identified needs of the organization as articulated by senior management.

Taking trainee skill or knowledge measurements during training programs: In a training program that is spread out over time or that requires the learning of specific sequential skills, it may be desirable to evaluate skill/knowledge gains during the course of training.

Following up after trainees have gone back to their jobs (3 to 12 months after training) to see if they are able to use their skill/knowledge on the job: Translating skills/knowledge learned in a workshop or trainer-controlled setting to the work environment is a critical step in adding to the worth of the training experience.

Monitoring training during the training activity to make course corrections: Being sure that the training design fits the trainees' needs often calls for the flexibility to make adjustments in mid-course.
Taking skill or knowledge measurements after training to determine trainee achievement: Being sure that the trainees leave the training situation with the desired knowledge/skill levels is part of being sure the training has merit.

Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training: One of the objectives of training is to increase productivity to the point at which the end product costs less with training than it does without.

Collecting data (rate of sales, number and type of customer complaints, employee turnover, rate of output, etc.) to provide evidence of training effect: Gathering production data before and after training, showing that training did make a difference, will help provide evidence of organizational effect.

Comparing trainee skill or knowledge measurements before and after training to determine gains: To measure actual gains in skill/knowledge, a baseline before the training event must be compared to the gains after the training event.

Soliciting judgments about training plans from participants: Testing ideas on the trainee population before the training event is one way to prevent a faulty training design.

Gathering data to compare training costs with the organizational impact of training: Adding up the real cost of the training event and comparing it with the actual gain in terms
of effect (which is profit for most organizations) will add to support for the worth of training.

4. That the respondents answered accurately and took the necessary time to respond thoughtfully to the survey questions.
CHAPTER III

METHODOLOGY AND PROCEDURES

The purpose of this chapter is to discuss the methods and procedures used to test the hypotheses. The chapter is divided into five sections: Development of the Instrument, Sampling Procedure, Data Collection, Treatment of the Data, and Analysis Procedures.

Development of the Instrument

The data for this study were collected from a systematic sample using a written questionnaire. A written questionnaire was used because it can be given to a large number of people in a fairly short time, it provides data that are easily tabulated and analyzed, this format is familiar to most respondents, subjects can respond when it is convenient for them, respondents have ample time to think about their responses, it permits anonymity, and using the questionnaire allows the researcher to cover a large geographic area.

A significant portion of the data gathered was perceptual data; that is, it was data based on the perceptions of the respondents. The possibility that the facts may not be the same as perceptions in any particular situation must be considered a limitation of this study.
Two questionnaires were developed for this study (Appendix A). One, Questionnaire B, was sent to the training director or person primarily in charge of training for the organization. The other, Questionnaire A, was sent to the person to whom the training director reported.

Part I of each questionnaire sought demographic data and the respondent's perception of training and evaluation in his/her organization. Part I of Questionnaire A asked for the name of the organization, the principal product or service offered, the type of organization, and the approximate number of employees in the organization. In addition, it asked if there was at least one unit in the organization whose major responsibility was training and to describe the reporting relationship of the training director through the chief executive officer or president. This part of Questionnaire A closed by asking the respondent to describe how important training was to the organization's success: extremely important, very important, neither important nor unimportant, not very important, or not at all important.

Part I of Questionnaire B contained questions designed to determine the scope of training activities: how many employees participated in training the past year and what percentage of training activities were conducted by the training staff, by external consultants, or considered to be on-the-job training. Six questions concerned the evaluation of training programs. Those questions asked what kind of training was evaluated and how often;
whether the respondent was satisfied with the quality of the training evaluation and, if not, why not; the motivation for training evaluation; who generally conducted the evaluation; with whom the results were shared; and whether upper management required that the evaluation results be shared with them.

Part II of both Questionnaires A and B listed 11 training-evaluation activities. These activities are intended to mirror the six-stage evaluation model proposed by Brinkerhoff (1987) and discussed in Chapters I and II. Stage I is reflected in Evaluation Activities 2 and 7, where tentative training goals are assessed in terms of their potential for worthwhile organizational benefit. Stage II is reflected in Evaluation Activity 10, where information is gathered to judge the training design. Stage III is reflected in Evaluation Activities 3 and 5, where the training design is monitored during implementation. Stage IV is reflected in Evaluation Activities 1, 6, and 9, where immediate trainee reactions and learning outcomes are assessed. Stage V is reflected in Evaluation Activity 4, where on-the-job use of training-acquired learning is measured. Stage VI is reflected in Evaluation Activities 8 and 11, where payoff to the organization is determined. Six of these activities (Activities 1, 3, 5, 6, 9, and 10) are aimed at judging the merit of training, and five of the activities (Activities 2, 4, 7, 8, and 11) are focused on evaluating the worth of training. For each of the 11 items, the respondents were asked
to judge whether or not that activity would (a) produce information that would generate management support for training, (b) provide information to help produce more effective training, and (c) be feasible in their organization and could be successfully carried out.

In addition, in Part II of Questionnaire B, the training directors were asked to judge how often a particular activity was carried out in their organizations. In making that estimate, the training directors were asked to indicate, on a 4-point scale, whether they usually, often, sometimes, or hardly ever used that activity for evaluation. A 4-point scale was used so that in the data analysis the responses could be collapsed into two categories: usually/often and sometimes/hardly ever.

This instrument was pilot tested by giving it to four upper-level managers and four training directors. These individuals were all in job situations similar to the population to be sampled (i.e., Foremost Corporation, Foremost Financial Services, Amway, and Steelcase). These persons were asked to respond to the questionnaire and to the tentative cover letter. Problems of question/answer wording, format of the questionnaire, internal validity of items, and the relationship among items were all evaluated in the pilot testing. To be useful, the questions and the answer categories needed to make sense to the respondents. A failure to answer, multiple answers when only one was asked for, too frequent "other" responses, and often-qualified answers were danger
signs. Both wording and format were examined when any of the above signs were seen. Looking at internal validation and the relationship among items allowed for the elimination of unnecessary questions or answers and in some cases the inclusion of other answer categories to gather more accurate data. Appropriate revisions and alterations in the format were made as a result of the pilot test.

Sampling Procedure

Three hundred twenty-three nonmanufacturing business organizations with 1,000 or more employees were selected using a systematic sampling procedure. The selections were made from the Dunn and Bradstreet (1986) Million Dollar Directory, which lists 3,554 nonmanufacturing businesses with 1,000 or more employees. Nonmanufacturing businesses with at least 1,000 employees were chosen for study because, according to results of a Bureau of National Affairs (1985) study on training, it is the large, nonmanufacturing business organizations that engage in the most training evaluation. The purpose of the study was not to discover what kinds of organizations perform training evaluations, but rather to determine, when such evaluation does occur, how it is done and some of the circumstances surrounding the evaluation. Therefore, the writer believed that aiming the study at the population most likely to do training evaluation would be most productive.
Data Collection

The first mailing for the study was sent on December 14, 1986. This mailing consisted of a letter and a return postcard and was sent to the president of each of the 323 companies in the sample (Appendix B). The company presidents were asked to indicate on the postcard the name of the person in the organization who was principally in charge of training and the name of the person to whom that training director reported. The initial mailing yielded responses from 59 companies or 18% of the sample. Forty-four of the returned cards contained the information requested, six individuals chose not to participate, and nine letters were returned as not deliverable. From early January to late February, the writer contacted all of the remaining 264 organizations by telephone in an effort to gain their participation in the study. Of that number, 106 agreed to participate and 158 chose not to do so. Of those choosing not to take part, 48 had policies against participating in research studies, 27 had gone out of business or had been purchased by other companies and were in the process of reorganizing, 17 were holding corporations for a number of smaller companies, and 66 had no identifiable training director or department.

Depending on how the company had been contacted (by telephone or letter) and whom the writer had contacted (training director, president, secretary), an appropriate letter and the questionnaire were sent on April 13, 1987 (Appendix C). A reply was requested by
April 27. This mailing yielded responses from 93 training directors and upper-level managers. Of that number, 56 were matched responses, that is, an A and a B survey from 28 organizations, and 37 were single responses, that is, either an A or a B survey from 37 organizations.

On May 8, another letter was sent to the persons from the 37 organizations who had not returned their surveys (Appendix D), and a postcard was sent to 85 upper-level managers and training directors who had not responded (Appendix E). This mailing yielded another 71 responses: 48 matched Survey A and B responses (24 companies), 20 single Survey A or B responses, and 3 refusals. On June 5, another attempt was made to solicit the missing Survey A or B responses (Appendix F), and another questionnaire was sent to the individuals in the 41 companies from whom there had been no response (Appendix G). This final mailing yielded an additional 64 Survey A and B responses (32 companies) and 38 single responses. The final total, then, was 203 responses: 83 organizations returned both Survey A and B (n = 166), 21 returned Survey B only, and 16 returned Survey A only. Twenty-four organizations did not respond to any of the mailings, and six said they would not participate in any way.

Information from the returned questionnaires was coded onto data sheets. The data were then entered into a computer data set.
Treatment of the Data

The independent variable in this study was the extent to which there was congruence in the perceptions of training directors and upper-level managers relative to the need for, value of, and feasibility of gathering complex and complete evaluation data for training, education, and development activities. These perceptions were judged using surveys given to training directors and upper-level managers in nonmanufacturing business organizations.

The dependent variable was the frequency with which training-program evaluation was accomplished. Frequency was judged on a 4-point scale (1 = usually [75-100% of the time], 2 = often [50-74% of the time], 3 = sometimes [25-49% of the time], and 4 = hardly ever [0-24% of the time]), using 11 evaluation activities drawn from the six-stage model for evaluation developed by Brinkerhoff (1987).

Analysis Procedures

The variables represented by each survey question were classified into different groups (nominal, ordinal, interval, ratio) according to how they were measured. The way in which the data were analyzed depended on how each variable was measured.

In analyzing Part I of the survey, if nominal data were gathered and the numerical code assigned to the possible responses conveyed no information, frequencies and percentages were used to indicate how upper-level managers and training directors responded. Examples
of nominal data included type of organization, who delivers training, satisfaction with training evaluation, impediments to training evaluation, who generally conducts training evaluation, with whom the training-evaluation results are shared, and whether management requires training evaluation.

If ordinal data were gathered on a scale ranging from little to much, the possible responses were arranged in order, giving both frequencies and percentages. Examples of ordinal data included categories indicating the number of employees participating in training, importance of training to the organization's success, frequency with which different types of training are evaluated, and the ranking of factors motivating the evaluation of training.

The number of employees in the organization represents ratio data. Respondents were asked to indicate the approximate number of employees in their organizations. The frequencies and means were given for these ratio data.

Part II of each survey listed 11 activities related to training evaluation. Upper-level managers and training directors were asked to decide whether each activity would help generate management support, would help produce more effective training, or would be feasible in their organizations. The response choices for each question were "yes" or "no," providing nominal data for this portion of Part II.

In addition, the training directors were asked if they usually (75-100% of the time), often (50-74% of the time), sometimes (25-49%
of the time), or hardly ever (0-24% of the time) accomplished each of these activities. The response choices provided ordinal data. In the statistical analysis, the categories were combined to produce only two values: usually/often and sometimes/hardly ever.

Frequencies and percentages of agreement between upper-level managers and training directors concerning whether or not each of the 11 evaluation activities would provide for management support, produce more effective training, or be feasible in their organizations were calculated. In addition, the mean frequency of training-evaluation activities was reported in cases in which upper-level managers' and training directors' responses indicated positive agreement, disagreement, or negative agreement.

Chi-square, lambda, Cramer's \( V \), and the \( t \)-test were the statistics used. These statistics were chosen because they were appropriate to the data collected and the hypotheses tested.

The survey methodology used produced primarily nominal and ordinal data. Given this level of measurement, the statistical measures used were appropriate, meaningful, and reliable. Other statistical measures, although more powerful and/or elegant, depend on certain assumptions that are violated by nominal or ordinal levels of measurement.

A chi-square test was used to measure the association between positive agreement, disagreement, and negative agreement on the part of upper-level managers and training directors concerning the value
of, need for, and feasibility of each of the 11 training-evaluation activities (independent variable) and the degree to which the training-evaluation activity was generally carried out (dependent variable).

In addition to the chi-square test, two measures of association for nominal data, Cramer’s V and lambda, were applied. Cramer’s V is a chi-square-based statistic and measures the strength of the association on a scale of 0 to 1. Consequently, high chi-square values and Cramer’s V’s approaching 1 will be positively related. Lambda reports the proportion by which the error can be reduced in predicting the dependent variable.

Finally, two other analyses using t-tests were performed to discover whether there was a difference in evaluation activity between organizations in which (a) training directors were satisfied with the quality of their training-evaluation activities and those in which training directors were not satisfied, and (b) training directors were required to report evaluation-activity results to upper-level management and those in which training directors were not required to report such results. For these analyses, the standardized score (t-value) and the degrees of freedom were used to determine the probability that the two variables were statistically independent.
CHAPTER IV

ANALYSIS AND DISCUSSION OF RESULTS

Introduction

This chapter presents an analysis and discussion of the data collected for the study and is divided into four sections. In the first section, the number of questionnaires returned and the types of respondents are reviewed. The second section contains a discussion of the characteristics of the population surveyed. The third gives the characteristics of evaluation activities as reported by training directors. Finally, the fourth section is a report of the data supplied by upper-level managers and training directors as they reported their perceptions of the value of, need for, and feasibility of 11 training-evaluation activities. In addition, the fourth section reveals the degree to which training directors reported accomplishing the 11 training-evaluation activities. The researcher attempted to identify findings of particular interest that may increase the understanding of factors that influence the accomplishment of complex training evaluation in an organizational setting.

Questionnaire Return

Questionnaires were mailed to 323 training directors and 323 upper-level managers representing 323 nonmanufacturing business
organizations with at least 1,000 employees. The business organizations were chosen from 3,554 nonmanufacturing businesses with at least 1,000 employees listed in Dunn and Bradstreet's (1986) Million Dollar Directory. Of the 323 businesses, 173 were eliminated because they did not have identifiable training directors, had gone out of business, or did not wish to participate in the study.

A total of 203 responses were received from the 150 companies that had training units and whose policies permitted their participation. One hundred sixty-six of those were matches representing 83 businesses. Sixteen companies' upper-level managers alone responded, and 21 firms' training directors alone responded. The data from every survey were usable. In reporting the upper-level managers' attitudes and responses, all 99 of their questionnaires were used. In reporting the training directors' attitudes and responses, all 104 of the questionnaires from these individuals were used. In comparing upper-level managers' responses with training directors' responses, only the 83 matched pairs were used.

Characteristics of the Population

The participants were asked to answer several demographic questions regarding their organization, such as the principal product/service offered, the type of organization, the number of employees in the organization, the approximate number of employees
participating in training, the reporting relationship of the training director, and a judgment of the importance of training to the organization's success (extremely important, very important, neither important nor unimportant, not very important, not at all important).

The upper-level managers and training directors were also asked to answer several questions about training evaluation. Those questions dealt with what types of training they evaluated; whether they were satisfied with their training evaluation; if not, what were the impediments to successful evaluation; how they would rank nine motivators to training evaluation; who generally conducted the evaluations; with whom the results were shared; and, if the results were shared with management, whether management required those results to be shared.

The response frequencies and percentages were calculated for each question, where appropriate. In some instances, not all subjects responded to all of the questions. An analysis of the data revealed the following information about the respondents' employing organizations.

**Type of Organization**

Upper-level managers were asked to check one of the following categories of organizations: merchandising or retail trade, service organization, utilities, entertainment, or other. One hundred
twenty usable responses were categorized, as shown in Table 1. The respondents who chose the category of "other" for the type of organization represented seven construction/engineering companies and one book-publishing firm.

Table 1

<table>
<thead>
<tr>
<th>Type of Organization</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merchandising/retail</td>
<td>27</td>
<td>22.5</td>
</tr>
<tr>
<td>Service organization</td>
<td>59</td>
<td>49.2</td>
</tr>
<tr>
<td>Utility</td>
<td>15</td>
<td>12.5</td>
</tr>
<tr>
<td>Entertainment</td>
<td>2</td>
<td>2.6</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>6.6</td>
</tr>
<tr>
<td>Missing response</td>
<td>9</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Number of Employees in the Organization

Upper-level managers were asked to indicate the approximate number of employees in their organization. The number of employees ranged from 1,000 to 60,000; the mean was 6,878. (See Table 2.)

Number of Employees Participating in Training

Concerning how many employees participated in training during the past year, training directors were asked to choose one of six
categories (0, 1-100, 100-300, 300-600, 600-900, 900 or more). The majority of respondents (36) chose the largest category: 900 or more. (See Table 3.)

Table 2
Number of Employees in the Organizations (N = 96; missing responses = 3)

<table>
<thead>
<tr>
<th>No. of Employees</th>
<th>No. of Respondents</th>
<th>No. of Employees</th>
<th>No. of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000-2,000</td>
<td>31</td>
<td>11,001-12,000</td>
<td>1</td>
</tr>
<tr>
<td>2,001-3,000</td>
<td>14</td>
<td>12,001-13,000</td>
<td>0</td>
</tr>
<tr>
<td>3,001-4,000</td>
<td>14</td>
<td>13,001-14,000</td>
<td>1</td>
</tr>
<tr>
<td>4,001-5,000</td>
<td>6</td>
<td>17,000</td>
<td>1</td>
</tr>
<tr>
<td>5,001-6,000</td>
<td>7</td>
<td>18,000</td>
<td>2</td>
</tr>
<tr>
<td>6,001-7,000</td>
<td>2</td>
<td>20,000</td>
<td>6</td>
</tr>
<tr>
<td>7,001-8,000</td>
<td>1</td>
<td>25,000</td>
<td>1</td>
</tr>
<tr>
<td>8,001-9,000</td>
<td>0</td>
<td>27,000</td>
<td>1</td>
</tr>
<tr>
<td>9,001-10,000</td>
<td>3</td>
<td>30,000</td>
<td>1</td>
</tr>
<tr>
<td>10,001-11,000</td>
<td>2</td>
<td>50,000</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60,000</td>
<td>1</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 3

Number of Employees Participating in Training (N = 103; missing response = 1)

<table>
<thead>
<tr>
<th>No. of Employees</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>1-100</td>
<td>7</td>
<td>6.8</td>
</tr>
<tr>
<td>100-300</td>
<td>24</td>
<td>23.3</td>
</tr>
<tr>
<td>300-600</td>
<td>22</td>
<td>21.4</td>
</tr>
<tr>
<td>600-900</td>
<td>14</td>
<td>13.6</td>
</tr>
<tr>
<td>900 or more</td>
<td>36</td>
<td>35.0</td>
</tr>
</tbody>
</table>

The Form Training Takes in the Organization

Training directors were asked what percentage of training in their organizations was delivered by the internal training unit, what percentage was provided by external consultants, and what percentage was informal, on-the-job training. The internal training units delivered by far the largest percentage of training (68.9%). (See Table 4.)

Importance of Training to the Organization’s Success

The upper-level managers were asked to rate the importance of training to the organization’s success. They were asked to check one of six choices: extremely important, very important, neither
important nor unimportant, not very important, or not at all important. The majority of respondents rated training as extremely to very important to the organization's success. (See Table 5.)

Table 4

Percentage of Training Provided by Various Groups (N = 104)

<table>
<thead>
<tr>
<th>Who Delivers</th>
<th>Mean Percentage of Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal training staff</td>
<td>68.9</td>
</tr>
<tr>
<td>Informal, on-the-job training</td>
<td>17.1</td>
</tr>
<tr>
<td>External consultants</td>
<td>13.9</td>
</tr>
</tbody>
</table>

Table 5

Importance of Training to the Organization's Success (N = 99)

<table>
<thead>
<tr>
<th>Importance</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely important</td>
<td>52</td>
<td>52.5</td>
</tr>
<tr>
<td>Very important</td>
<td>39</td>
<td>39.4</td>
</tr>
<tr>
<td>Neither important nor unimportant</td>
<td>6</td>
<td>6.1</td>
</tr>
<tr>
<td>Not very important</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>Not at all important</td>
<td>1</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Characteristics of Evaluation Activities, as Reported by Training Directors

One hundred four training directors responded to seven questions concerning training evaluation. Their answers are discussed on the following pages.

What Kind of Training Is Evaluated and How Often?

One hundred two training directors responded to the question concerning whether training conducted by the internal training staff was evaluated. They were asked to check one of three responses: yes, always; yes, sometimes; or no, never. The frequencies and percentages of responses are reported in Table 6.

Table 6
Evaluation of Internal Training (N = 102; missing responses = 2)

<table>
<thead>
<tr>
<th>Question: Do you evaluate training conducted by you or your staff?</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, always</td>
<td>75</td>
<td>73.5</td>
</tr>
<tr>
<td>Yes, sometimes</td>
<td>24</td>
<td>23.5</td>
</tr>
<tr>
<td>No, never</td>
<td>3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Ninety-one training directors responded to the question regarding whether training conducted by external consultants was
evaluated. They were asked to check one of three responses: yes, always; yes, sometimes; or no, never. The frequencies and percentages of responses are reported in Table 7.

Table 7

<table>
<thead>
<tr>
<th>Question: Do you evaluate training conducted by external consultants?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Yes, always</td>
</tr>
<tr>
<td>Yes, sometimes</td>
</tr>
<tr>
<td>No, never</td>
</tr>
</tbody>
</table>

Ninety-seven training directors responded to the item concerning whether informal, on-the-job training was evaluated. They were asked to check one of three responses: yes, always; yes, sometimes; or no, never. Response frequencies and percentages are reported in Table 8.

Satisfaction With Training Evaluation

Participants were asked if they were satisfied with the quality of the training program evaluation they performed. They were asked to check one of three responses: yes; no; or N/A, we don't
evaluate. If the participant indicated that she/he did not evaluate, the individual was asked to go directly to Part II of the survey. Slightly more respondents said they were satisfied than said they were not. The exact figures are given in Table 9.

Table 8
Evaluation of On-The-Job Training (N = 97; missing responses = 7)

<table>
<thead>
<tr>
<th>Response</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, always</td>
<td>13</td>
<td>13.4</td>
</tr>
<tr>
<td>Yes, sometimes</td>
<td>50</td>
<td>51.5</td>
</tr>
<tr>
<td>No, never</td>
<td>34</td>
<td>35.1</td>
</tr>
</tbody>
</table>

Table 9
Satisfaction With the Quality of Training Evaluation (N = 104)

<table>
<thead>
<tr>
<th>Response</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>53</td>
<td>51.0</td>
</tr>
<tr>
<td>No</td>
<td>48</td>
<td>46.2</td>
</tr>
<tr>
<td>N/A, we don’t evaluate</td>
<td>3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Respondents who were not satisfied with the quality of their training program evaluations (n = 48) were asked to indicate all
impediments that might be affecting their evaluation activities. To do this, they were to check all of the following items that applied: lack of budget, lack of time, lack of expertise, lack of evaluation methods, lack of staff, not required by a higher authority, not considered important by the training director, and "other." Table 10 shows the number and percentage of respondents who checked particular items as impediments.

Table 10
Percentage of Respondents Reporting Particular Items as Impediments to Quality Training Evaluation (N = 48)

<table>
<thead>
<tr>
<th>Impediment</th>
<th>No. of Respondents</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of time</td>
<td>40</td>
<td>83.3</td>
</tr>
<tr>
<td>Lack of staff</td>
<td>29</td>
<td>60.4</td>
</tr>
<tr>
<td>Lack of adequate evaluation methods</td>
<td>24</td>
<td>50.0</td>
</tr>
<tr>
<td>Not required by a higher authority</td>
<td>19</td>
<td>39.5</td>
</tr>
<tr>
<td>Lack of budget</td>
<td>17</td>
<td>35.4</td>
</tr>
<tr>
<td>Lack of expertise</td>
<td>14</td>
<td>29.1</td>
</tr>
<tr>
<td>Not considered important by training manager</td>
<td>3</td>
<td>6.2</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>10.4</td>
</tr>
</tbody>
</table>
Motivation to Accomplish Training Evaluation

Training directors who reported performing some evaluation activity were asked to rank eight factors in the order of their importance in motivating them to evaluate training. A ranking of 1 indicated the highest, most motivating factor. The mean rankings, from highest to lowest, given by the 100 training directors who responded to this question are shown in Table 11.

Table 11

Mean Ranking of Factors Motivating Training Directors to Evaluate Training (N = 100; missing responses = 4)

<table>
<thead>
<tr>
<th>Motivating Factor</th>
<th>Mean of Rank Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used to determine the effectiveness of training</td>
<td>2.00</td>
</tr>
<tr>
<td>Used to improve the training design for future programs</td>
<td>2.69</td>
</tr>
<tr>
<td>Used to measure the application of training back on the job</td>
<td>3.88</td>
</tr>
<tr>
<td>Used to measure progress toward the company’s objectives/profit</td>
<td>4.64</td>
</tr>
<tr>
<td>Used to provide feedback for trainees on their progress in knowledge/skills</td>
<td>4.67</td>
</tr>
<tr>
<td>Used to justify training programs</td>
<td>5.21</td>
</tr>
<tr>
<td>Used to provide evidence of training’s cost effectiveness</td>
<td>5.69</td>
</tr>
<tr>
<td>Required by a higher authority</td>
<td>6.82</td>
</tr>
</tbody>
</table>
The training directors were divided into two groups: those whose managers required them to share evaluation results ($n = 32$) and those whose managers did not require them to share such results ($n = 68$). The responses of each group to the ranking of motivators, in order of importance, are given in Table 12.

Who Conducts Training Evaluation

Training directors were asked who conducted evaluation for the three kinds of training (conducted by training staff, conducted by external consultants, or on-the-job training). The choices were: the training staff, an ad hoc committee, in-house evaluation experts, an outside consultant, supervisors, or others. The respondents were asked to check all choices that applied. For example, a respondent might check both the training staff and supervisors if both had a role in evaluating internal training. In every case in which the category "others" was chosen, the respondents indicated that it was the trainees or participants in the training who conducted the evaluation. (See Table 13.)

With Whom the Evaluation Results Were Shared

Training directors were asked to indicate with whom they shared their evaluation results. Table 14 shows the percentage of respondents who indicated various categories of individuals with whom they shared the evaluation results. Respondents were
Table 12
Ranking of Motivating Factors in Evaluating Training by Those Required (n = 32) and Not Required (n = 68) to Share Results (missing responses = 4)

<table>
<thead>
<tr>
<th>Motivating Factor</th>
<th>Mean Rank by Those Required to Share</th>
<th>Rank</th>
<th>Mean Rank by Those Not Required to Share</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used to determine the effectiveness of training</td>
<td>2.03</td>
<td>1</td>
<td>2.00</td>
<td>1</td>
</tr>
<tr>
<td>Used to improve the training design for future programs</td>
<td>2.76</td>
<td>2</td>
<td>2.67</td>
<td>2</td>
</tr>
<tr>
<td>Used to measure the application of training back on the job</td>
<td>3.97</td>
<td>3</td>
<td>3.83</td>
<td>3</td>
</tr>
<tr>
<td>Used to measure progress toward the company's objectives/profit</td>
<td>4.72</td>
<td>4</td>
<td>4.59</td>
<td>5</td>
</tr>
<tr>
<td>Used to justify training programs</td>
<td>5.08</td>
<td>5</td>
<td>5.24</td>
<td>6</td>
</tr>
<tr>
<td>Used to provide evidence of training's cost effectiveness</td>
<td>5.56</td>
<td>6</td>
<td>5.75</td>
<td>7</td>
</tr>
<tr>
<td>Used to provide feedback for trainees on their progress in knowledge/skills</td>
<td>5.70</td>
<td>7</td>
<td>4.22</td>
<td>4</td>
</tr>
<tr>
<td>Required by a higher authority</td>
<td>5.96</td>
<td>8</td>
<td>7.24</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 13
Percentage of Training Directors Who Identified the Following as Conducting Training-Evaluation Activities (n = 101; missing responses = 3)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Training staff</td>
<td>89.1</td>
<td>62.4</td>
<td>31.7</td>
</tr>
<tr>
<td>Supervisors</td>
<td>51.5</td>
<td>27.7</td>
<td>50.5</td>
</tr>
<tr>
<td>In-house experts</td>
<td>10.9</td>
<td>9.9</td>
<td>0.0</td>
</tr>
<tr>
<td>Ad hoc committee</td>
<td>9.9</td>
<td>8.9</td>
<td>2.0</td>
</tr>
<tr>
<td>Outside consultants</td>
<td>4.0</td>
<td>6.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Others (trainees)</td>
<td>10.9</td>
<td>7.9</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Table 14
Percentage of Training Directors Who Identified Persons With Whom Training-Evaluation Results Were Shared (n = 101; missing responses = 3)

<table>
<thead>
<tr>
<th>Specified Persons</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Others in training department</td>
<td>87.1</td>
</tr>
<tr>
<td>The person to whom I report</td>
<td>87.1</td>
</tr>
<tr>
<td>The person conducting the training</td>
<td>82.2</td>
</tr>
<tr>
<td>Supervisor or manager of trainee</td>
<td>66.3</td>
</tr>
<tr>
<td>Supervisor requesting training</td>
<td>41.6</td>
</tr>
<tr>
<td>Trainees</td>
<td>27.7</td>
</tr>
<tr>
<td>Others (senior management)</td>
<td>15.8</td>
</tr>
</tbody>
</table>
instructed to check all that applied. Those who specified "others" all indicated that they shared results with members of senior management.

Management Requiring Training Evaluations

Training directors were asked, "If you share training evaluation results with anyone in management above you, do they require it?" One hundred training directors responded, of whom 32 said they were required to share evaluation results and 68 said they were not required to share such results.

Perceptions of Upper-Level Managers and Training Directors Concerning the Value of, Need for, and Feasibility of Training-Evaluation Activities and the Degree to Which the Activities Were Carried Out

Eleven activities related to training evaluation were listed in Part II of the survey. The upper-level managers and training directors were asked to decide independently if each activity would help generate management support (value), help produce more effective training (need), or be feasible in their organization. In addition, the training directors were asked if they usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), or hardly ever (0-24% of the time) accomplished each of those activities.

This section of Chapter IV reports the results of the upper-level managers' and training directors' perceptions of
training-evaluation activities, as well as the frequency with which those activities were accomplished, as reported by the training directors. All 99 surveys completed by upper-level managers were used in reporting their responses. All 104 surveys completed by training directors were used in reporting their responses. In comparing upper-level managers' and training directors' responses, the 83 matched pairs (total 166 surveys) were used.

Because the measurements used in Part II employed nominal and ordinal data, the analyses of the data were limited to those the measurement would allow. Statistical tests used in making decisions about the research hypotheses were chosen for (a) the power of the tests, (b) the application of the statistical model on which the test was based to the research data, and (c) the level of measurement achieved in the research.

Agreement Between Upper-Level Managers and Training Directors Concerning Evaluation Activities and the Frequency With Which the Activities Were Carried Out

For each of the 11 evaluation activities, the training directors \((n = 104)\) were asked to indicate whether they usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), or hardly ever (0-24% of the time) accomplished those activities. The responses were collapsed so that "usually" and "often" were one category and "sometimes" and "hardly ever" were another category. In addition, each item was labeled to reflect
whether it measured the merit or worth of training. Those labels reflected the Six-Stage Evaluation Model formulated by Brinkerhoff (1987). See Chapter III, page 52, for a complete explanation of how the activity labels were determined. The evaluation activities, listed in order of the frequency with which they were carried out, are shown in Table 15, Column 1.

The most frequently conducted training-evaluation activity was the satisfaction rating. With one exception (formulating training outcomes to reflect organizational need), all merit evaluation activities were carried out more frequently by training directors than were worth activities.

For each of the evaluation activities, the upper-level managers and training directors were asked to indicate whether or not they thought that activity would help generate management support, help produce more effective training, or be feasible in their organizations.

Columns 2, 3, and 4 of Table 15 show the percentage of all training directors and upper-level managers (N = 203) who agreed that the particular evaluation activity would be feasible in their organization, would help produce more effective training, or would help generate management support. The numbers in parentheses indicate rank order within the columns.

It is interesting that respondents agreed that it is the evaluation activities demonstrating worth that are most likely to
Table 15
Frequency With Which Evaluation Activities Were Reported to Be Generally Carried Out and the Percentage of Upper-Level Managers (ULM) and Training Directors (TD) Who Agreed That the Various Activities Were Feasible, Would Produce More Effective Training, and Would Help Generate Management Support

<table>
<thead>
<tr>
<th>Merit or Worth?</th>
<th>Evaluation Activity</th>
<th>Percent of TDs Who Reported Generally Carrying Out the Activity (n=104)</th>
<th>Percent of ULMs &amp; TDs Who Felt the Activity Was Feasible (n=203)</th>
<th>Percent of ULMs &amp; TDs Who Felt the Activity Would Produce More Effective Training (n=203)</th>
<th>Percent of ULMs &amp; TDs Who Felt the Activity Would Help Generate Management Support (n=203)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merit</td>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>85.5 (1) 8</td>
<td>96.6 (1) 9</td>
<td>88.2 (6) 7</td>
<td>73.9 (7) 4</td>
</tr>
<tr>
<td>Merit</td>
<td>Monitoring training during the training activity to make course corrections</td>
<td>59.0 (2) 5</td>
<td>84.2 (4) 2</td>
<td>92.1 (2) 2</td>
<td>45.3 (10) 3</td>
</tr>
<tr>
<td>Worth</td>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>37.6 (3) 1</td>
<td>86.7 (2) 0</td>
<td>70.6 (3) 0</td>
<td>88.7 (1) 4</td>
</tr>
<tr>
<td>Merit</td>
<td>Soliciting judgments about training plans from participants</td>
<td>55.4 (4) 1</td>
<td>85.7 (3) 0</td>
<td>86.2 (8) 0</td>
<td>52.2 (9) 0</td>
</tr>
<tr>
<td>Merit</td>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>42.2 (5) 1</td>
<td>74.4 (7) 0</td>
<td>88.6 (5) 0</td>
<td>43.8 (11) 0</td>
</tr>
<tr>
<td>Merit</td>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>34.9 (6) 1</td>
<td>73.9 (8) 0</td>
<td>67.2 (7) 0</td>
<td>67.5 (8) 0</td>
</tr>
<tr>
<td>Merit</td>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>32.5 (7) 1</td>
<td>75.9 (6) 0</td>
<td>89.2 (4) 0</td>
<td>76.8 (6) 0</td>
</tr>
<tr>
<td>Worth</td>
<td>Following up after trainees have gone back to their jobs (3-12 mos. after training) to see if they are able to use their skills/knowledge on the job</td>
<td>30.1 (8) 1</td>
<td>81.3 (5) 0</td>
<td>94.6 (1) 0</td>
<td>81.7 (4) 0</td>
</tr>
<tr>
<td>Worth</td>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>24.1 (9) 1</td>
<td>69.9 (9) 0</td>
<td>51.2 (11) 0</td>
<td>04.2 (2) 0</td>
</tr>
<tr>
<td>Worth</td>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>10.1 (10) 1</td>
<td>54.2 (11) 0</td>
<td>69.5 (9) 0</td>
<td>78.3 (5) 0</td>
</tr>
<tr>
<td>Worth</td>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>10.1 (11) 1</td>
<td>60.1 (10) 0</td>
<td>57.6 (10) 0</td>
<td>82.3 (3) 0</td>
</tr>
</tbody>
</table>

*Numbers in parentheses indicate rank order within the column.*
generate management support (Table 15, Column 4). Producing more effective training, respondents agreed, is more likely to be a function of activities demonstrating merit (Table 15, Column 3). Training-evaluation activities determining the merit of training are most likely, according to rank, to be feasible as well (Table 15, Columns 2 and 3). Tying the results of training to the bottom line was perceived as the least feasible training-evaluation activity (Table 15, Column 2) and also the least likely to be carried out (Table 15, Column 1).

Specifically formulating training outcomes to reflect organizational need was an activity that 88.7% of the respondents thought would help generate management support (Table 15, Column 4), 86.7% thought it was feasible (Table 15, Column 2), and yet it was generally carried out by only 57.8% of the training directors (Table 15, Column 1). Following up after trainees have gone back to their jobs (3 to 12 months after training) to see if they are able to use their skills/knowledge on the job was an activity that 91.6% of the respondents thought would produce more effective training (Table 15, Column 3), 81.3% thought it was feasible (Table 15, Column 2), and yet only 30.1% of the training directors reported generally carrying out this evaluation activity. The most feasible activity as perceived by upper-level managers and training directors is end-of-training evaluation concerning trainee satisfaction with the training experience (96.6%--Table 15, Column 2). This is also the training-evaluation activity most frequently carried out by training
directors. Judging by comparing rankings, the closest correlation existed between the percentage of respondents who agreed about the feasibility of the activity (Table 15, Column 2) and the percentage of respondents who reported generally carrying out the activity (Table 15, Column 1).

**Frequency of Training-Evaluation Activity For Groups Reporting Agreement and Disagreement**

The amount of agreement/disagreement on the value of, need for, and feasibility of each evaluation activity is shown in Table 16 by giving the numbers of pairs of upper-level managers and training directors reporting positive agreement (yes/yes), disagreement (yes/no or no/yes), or negative agreement (no/no). In addition, the mean frequency with which each activity was carried out is shown for each group of responses. The group mean was calculated by using the following categories: usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), and hardly ever (0-24% of the time). Each category was assigned a value as follows: usually = 1, often = 2, sometimes = 3, and hardly ever = 4.

Of the 33 items listed, 82% showed more frequent activity when the condition of positive agreement existed. The activities that occurred most often, with the exception of the second item, were all items evaluating the merit of training. Consistently falling in the "usually" and "often" categories were Items 1, 2, 3, 5, and 10. In
Table 16

Frequency of Training-Evaluation Activity, Reported by the Means for Three Groups: Those Reporting Positive Agreement, Those Disagreeing, and Those Reporting Negative Agreement (N = 166-83 Pairs)

<table>
<thead>
<tr>
<th>Merit or Worth?</th>
<th>Evaluation Activity</th>
<th>Means for:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive Agreement</td>
</tr>
<tr>
<td>Merit</td>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>1.42 (50)^a</td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>1.39 (69)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>1.42 (79)</td>
</tr>
<tr>
<td>Worth</td>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>2.14 (69)</td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.17 (72)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.08 (69)</td>
</tr>
<tr>
<td>Merit</td>
<td>Taking trainee skill or knowledge measurements during the training program</td>
<td>2.67 (22)</td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.47 (65)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.24 (49)</td>
</tr>
</tbody>
</table>

^a Mean significance level.
## Table 16: Continued

<table>
<thead>
<tr>
<th>Merit or Worth?</th>
<th>Evaluation Activity</th>
<th>Means for:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive Agreement</td>
<td>Disagreement</td>
<td>Negative Agreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Worth</td>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their knowledge/skills on the job</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.85 (60)</td>
<td>3.33 (18)</td>
<td>3.00 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.91 (76)</td>
<td>3.40 (5)</td>
<td>4.00 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.61 (57)</td>
<td>3.85 (20)</td>
<td>3.60 (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merit</td>
<td>Monitoring training during the training activity to make course corrections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.26 (24)</td>
<td>2.39 (28)</td>
<td>2.48 (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.26 (72)</td>
<td>3.40 (10)</td>
<td>3.00 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.20 (62)</td>
<td>2.81 (16)</td>
<td>3.60 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merit</td>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.47 (43)</td>
<td>3.25 (28)</td>
<td>3.27 (12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.75 (68)</td>
<td>3.40 (10)</td>
<td>2.60 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.53 (56)</td>
<td>3.41 (17)</td>
<td>3.70 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worth</td>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.95 (61)</td>
<td>3.53 (18)</td>
<td>3.50 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>3.04 (27)</td>
<td>3.11 (28)</td>
<td>3.14 (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.69 (46)</td>
<td>3.59 (23)</td>
<td>3.69 (14)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 16: Continued

<table>
<thead>
<tr>
<th>Merit or Worth?</th>
<th>Evaluation Activity</th>
<th>Means for:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive Agreement</td>
</tr>
<tr>
<td>Worth</td>
<td>Collecting data (rate of sales, number and type of customer complaints, etc.) to provide evidence of training impact</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>3.28 (54)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>3.02 (42)</td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.69 (28)</td>
</tr>
<tr>
<td>Merit</td>
<td>Compare trainee skill or knowledge measurements before and after training to determine gains</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>3.00 (52)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.84 (69)</td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.71 (53)</td>
</tr>
<tr>
<td>Merit</td>
<td>Soliciting judgments about training plans from participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>2.16 (33)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>2.20 (65)</td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>2.11 (67)</td>
</tr>
<tr>
<td>Worth</td>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- will help generate management support</td>
<td>3.25 (58)</td>
</tr>
<tr>
<td></td>
<td>- will help produce more effective training</td>
<td>3.39 (32)</td>
</tr>
<tr>
<td></td>
<td>- would be feasible in their organization</td>
<td>3.23 (36)</td>
</tr>
</tbody>
</table>

Numbers in parentheses indicate pairs of respondents.
18 of 33 cases, or 55% of the time, even when positive agreement existed the evaluation activities were generally not occurring. In 12 of those 18 cases, the evaluation activity was measuring the worth of training.

The Association Between Positive Agreement, Disagreement, and Negative Agreement on the Part of Upper-Level Managers and Training Directors and the Degree to Which the Evaluation Activities Were Generally Carried Out

A chi-square test was used to measure the relationship of the independent variable to the dependent variable. The independent variable was the extent to which there was positive congruence in the perceptions of upper-level managers and training directors relative to the value of (would generate management support), need for (would produce more effective training), and feasibility of (feasible in their organizations) gathering complete and complex training-evaluation data. The dependent variable was the degree to which the evaluation activity was generally carried out. Given 3 degrees of freedom, a chi-square score of 7.815 or higher with a significance of .05 or lower was considered to show a relationship.

The independent variable was divided into three parts: (a) upper-level managers and training directors who showed positive agreement on the value of, need for, or feasibility of a particular training-evaluation activity; (b) upper-level managers and training directors who disagreed on the value of, need for, or feasibility of
a particular training-evaluation activity; and (c) upper-level managers and training directors who showed negative agreement on the value of, need for, or feasibility of a particular training-evaluation activity.

To gather information on the dependent variable, the training directors were asked to indicate whether each of the 11 training-evaluation activities was carried out in their organizations usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), or hardly ever (0-24% of the time). For purposes of the chi-square test, these four categories were collapsed into two: usually/often and sometimes/hardly ever.

In addition to the chi-square test, two measures of association for nominal data, Cramer's $V$ and lambda, were used. Cramer's $V$ is a chi-square-based statistic that measures the strength of the association on a scale of 0 to 1. Consequently, high chi-square values and Cramer's $V$'s approaching 1 will be positively related. Lambda indicates the percentage proportion by which error can be reduced in predicting the dependent variable if one already knows the value of the independent variable. Only the 166 surveys representing 83 matched pairs (upper-level managers and training directors in the same company) were used in these analyses.

Table 17 shows the chi-square values, significance, Cramer's $V$, and lambda for responses concerning the frequency with which each of the 11 training-evaluation activities was carried out and respondents' perceptions of whether or not those activities would
Table 17

Association Between Agreement or Disagreement About the Value of the Training-Evaluation Activity to Help Produce Management Support and the Frequency With Which the Activity Is Carried Out (N = 166--83 pairs)

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Chi-Square</th>
<th>Significance*</th>
<th>Cramer's V</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>0.056</td>
<td>.972</td>
<td>0.026</td>
<td>0.000</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>3.526</td>
<td>.172</td>
<td>0.209</td>
<td>0.087</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>0.633</td>
<td>.729</td>
<td>0.011</td>
<td>0.088</td>
</tr>
<tr>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their skills/knowledge on the job</td>
<td>2.074</td>
<td>.355</td>
<td>0.159</td>
<td>0.000</td>
</tr>
<tr>
<td>Monitoring training during the training activity to make course corrections</td>
<td>0.096</td>
<td>.953</td>
<td>0.034</td>
<td>0.000</td>
</tr>
<tr>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>6.459</td>
<td>.039*</td>
<td>0.281</td>
<td>0.000</td>
</tr>
<tr>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>3.262</td>
<td>.163</td>
<td>0.210</td>
<td>0.000</td>
</tr>
<tr>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>0.239</td>
<td>.887</td>
<td>0.053</td>
<td>0.000</td>
</tr>
<tr>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>0.959</td>
<td>.619</td>
<td>0.107</td>
<td>0.000</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td>2.361</td>
<td>.307</td>
<td>0.171</td>
<td>0.056</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>2.727</td>
<td>.256</td>
<td>0.182</td>
<td>0.000</td>
</tr>
</tbody>
</table>

*aCells with expected frequencies less than 5 were too great to draw conclusions.

*Probability that the two variables are statistically independent.
help generate management support (value). Only one activity, taking
skill or knowledge measurements after training to determine trainee
achievement, showed a 96% probability that there was a relationship
between disagreement and agreement on the value of this activity to
produce management support and the frequency with which the activity
was carried out. The Cramer's V indicated the relationship was
rather weak. The lambda of .000 showed that one could not reduce
the error in predicting the frequency of activity by knowing whether
or not agreement or disagreement existed concerning the value of the
evaluation activity. Therefore, the results, as shown in Table 17,
do not support directional Hypothesis 1.

Table 18 contains the chi-square values, significance, Cramer's
V, and lambda for responses concerning the frequency with which each
of the 11 training-evaluation activities was carried out and
respondents' perceptions of whether or not those activities would
help produce more effective training (need). Statistics indicated
that there were three activities (the third, the fifth, and the
eighth) where an association might exist between agreement or
disagreement on the need for that activity to help produce more
effective training and the frequency with which the activity was
carried out. In the third and fifth activities, however, more than
20% of the cells had an expected frequency of less than five. This
condition prevented drawing any conclusions about these
relationships. In the eighth activity there was a 96% probability
that agreement or disagreement on the need for this activity was
Table 18

Association Between Agreement or Disagreement About the Need for Training-Evaluation Activity to Help Produce More Effective Training and the Frequency With Which the Activity Is Carried Out (N = 166-83 pairs)

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Chi-Square</th>
<th>Significance*</th>
<th>Cramer’s ¥</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>2.770</td>
<td>.250</td>
<td>0.026</td>
<td>0.000</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>3.356</td>
<td>.187</td>
<td>0.204</td>
<td>0.088</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>6.100</td>
<td>.047</td>
<td>0.237</td>
<td>0.000</td>
</tr>
<tr>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their skills/knowledge on the job</td>
<td>1.182</td>
<td>.554</td>
<td>0.119</td>
<td>0.000</td>
</tr>
<tr>
<td>Monitoring training during the training activity to make course corrections</td>
<td>7.696</td>
<td>.021</td>
<td>0.305</td>
<td>0.133</td>
</tr>
<tr>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>1.138</td>
<td>.566</td>
<td>0.117</td>
<td>0.000</td>
</tr>
<tr>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>0.202</td>
<td>.904</td>
<td>0.049</td>
<td>0.000</td>
</tr>
<tr>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>6.394</td>
<td>.041*</td>
<td>0.277</td>
<td>0.000</td>
</tr>
<tr>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>0.978</td>
<td>.613</td>
<td>0.109</td>
<td>0.000</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td>4.441</td>
<td>.109</td>
<td>0.233</td>
<td>0.139</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>1.189</td>
<td>.552</td>
<td>0.121</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* Cells with expected frequencies less than 5 were too great to draw conclusions.

*Probability that the two variables are statistically independent.
related to the frequency with which the activity was carried out. The Cramer's $V$ showed a relatively weak association. The lambda indicated one could not increase the accuracy of predicting whether the activity generally occurred by knowing whether there was agreement or disagreement on the need for the evaluation activity to help produce more effective training. Therefore, the results, as shown in Table 18, do not support directional Hypothesis 2.

Table 19 shows the chi-square values, significance, Cramer's $V$, and lambda for responses concerning the frequency with which each of the 11 training-evaluation activities was carried out and respondents' perceptions of whether or not those activities would be feasible in their organizations. Statistical analysis for the association between agreement and disagreement concerning the feasibility of carrying out training-evaluation activities and the frequency with which the activity was carried out showed a relationship in 7 of 11 cases. Of those seven, the third, the sixth, and the seventh had expected frequencies in all of the cells large enough to allow the writer to draw some conclusions. All three showed a very strong probability that there was an association in the population. The Cramer's $V$ for all three indicated a moderate relationship. The lambda for the third activity indicated that the accuracy of predicting the frequency for this evaluation activity could be increased by 15% if one knew whether there was agreement or disagreement between the upper-level manager and the training director on this item.
Table 19
Association Between Agreement or Disagreement About the Feasibility to Carry Out the Training-Evaluation Activity in the Organization and the Frequency With Which the Activity Is Carried Out (N = 166-83 pairs)

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Chi-Square</th>
<th>Significance*</th>
<th>Cramer’s V</th>
<th>Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>12.085</td>
<td>.002*</td>
<td>0.384</td>
<td>0.133</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>6.182</td>
<td>.045*</td>
<td>0.280</td>
<td>0.182</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>10.898</td>
<td>.004*</td>
<td>0.367</td>
<td>0.152</td>
</tr>
<tr>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their skills/knowledge on the job</td>
<td>14.882</td>
<td>.000*</td>
<td>0.426</td>
<td>0.000</td>
</tr>
<tr>
<td>Monitoring training during the training activity to make course corrections</td>
<td>7.069</td>
<td>.019*</td>
<td>0.309</td>
<td>0.206</td>
</tr>
<tr>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>8.240</td>
<td>.016*</td>
<td>0.317</td>
<td>0.000</td>
</tr>
<tr>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>12.703</td>
<td>.001*</td>
<td>0.398</td>
<td>0.000</td>
</tr>
<tr>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>5.265</td>
<td>.071*</td>
<td>0.253</td>
<td>0.000</td>
</tr>
<tr>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>3.418</td>
<td>.181*</td>
<td>0.204</td>
<td>0.000</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td>10.199</td>
<td>.006*</td>
<td>0.356</td>
<td>0.257</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>5.780</td>
<td>.055*</td>
<td>0.269</td>
<td>0.000</td>
</tr>
</tbody>
</table>

*Cells with expected frequencies less than 5 were too great to draw conclusions.
*Probability that the two variables are statistically independent.
Four of the seven had 20% or more cells with expected frequencies less than 5. Definite conclusions cannot be drawn about these cases even though it may very well be the case that the two variables are associated. The Cramer's V's of over .30 show the relationship of four of the seven to be moderately strong, and four of the lambda statistics show an increase in the possibility of predicting the dependent variable (the frequency of evaluation activity) if one knows the independent variable (agreement or disagreement on feasibility).

In addition, Table 16 showed more frequency in the evaluation-activity means related to positive agreement on feasibility than the means related to disagreement or negative agreement on feasibility. Where there is agreement on feasibility, 5 of 11 evaluation activities are generally carried out. The results of the study do not conclusively support Hypothesis 3, but they do indicate it may be perceived feasibility rather than need or value that is associated with the frequency of evaluation activities.

Disagreement Between Upper-Level Managers and Training Directors and How It Affects Frequency of Evaluation Activities

Table 20 contains information about disagreement in responses and how that may relate to whether or not the evaluation activity was accomplished. The means for the frequency of evaluation activities were computed for situations in which the upper-level manager and the training director disagreed in their responses. For
Table 20

Mean Scores of Actual Evaluation Activity When Upper-Level Managers (ULM) and Training Directors (TD) Disagreed on Whether the Activity Would Generate Management Support, Help Produce Effective Training or Would Be Feasible in Their Organization

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>t-Value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>ULM: Yes (8) TD: No (13)</td>
<td>1.38</td>
<td>11.8</td>
<td>.459</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>ULM: No (7) TD: Yes (3)</td>
<td>2.43</td>
<td>6.0</td>
<td>.047*</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>ULM: Yes (1) TD: No (2)</td>
<td>4.00</td>
<td>0.0</td>
<td>1.000</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>ULM: Yes (6) TD: No (5)</td>
<td>4.00</td>
<td>3.0</td>
<td>.016*</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>ULM: No (6) TD: Yes (0)</td>
<td>3.20</td>
<td>4.0</td>
<td>.005*</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>ULM: Yes (7) TD: No (3)</td>
<td>3.57</td>
<td>2.4</td>
<td>.252</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training program</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>ULM: Yes (11) TD: No (21)</td>
<td>2.13</td>
<td>30.0</td>
<td>.081</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>ULM: No (8) TD: Yes (8)</td>
<td>2.88</td>
<td>8.4</td>
<td>.042*</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>ULM: Yes (13) TD: No (11)</td>
<td>2.91</td>
<td>19.5</td>
<td>.050*</td>
</tr>
</tbody>
</table>
Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their knowledge/skills on the job

- will help generate management support
- will help produce more effective training
- would be feasible in their organization

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>ULM: Yes</th>
<th>ULM: No</th>
<th>TD: Yes</th>
<th>TD: No</th>
<th>t-value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3.86 (7)</td>
<td>3.00 (11)</td>
<td>2.23</td>
<td>12.9</td>
<td>.044*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.50 (2)</td>
<td>3.33 (3)</td>
<td>0.20</td>
<td>3.0</td>
<td>.854</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.00 (15)</td>
<td>3.40 (5)</td>
<td>2.45</td>
<td>4.0</td>
<td>.070</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Monitoring training during the training activity to make course corrections

- will help generate management support
- will help produce more effective training
- would be feasible in their organization

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>ULM: Yes</th>
<th>ULM: No</th>
<th>TD: Yes</th>
<th>TD: No</th>
<th>t-value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.58 (19)</td>
<td>2.00 (9)</td>
<td>1.19</td>
<td>15.2</td>
<td>.239</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.00 (6)</td>
<td>2.50 (4)</td>
<td>1.73</td>
<td>3.0</td>
<td>.182</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.86 (7)</td>
<td>2.00 (9)</td>
<td>4.01</td>
<td>9.6</td>
<td>.003*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Taking skill or knowledge measurements after training to determine trainee achievement

- will help generate management support
- will help produce more effective training
- would be feasible in their organization

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>ULM: Yes</th>
<th>ULM: No</th>
<th>TD: Yes</th>
<th>TD: No</th>
<th>t-value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3.38 (16)</td>
<td>3.08 (12)</td>
<td>0.76</td>
<td>24.2</td>
<td>.456</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.00 (6)</td>
<td>2.50 (4)</td>
<td>2.32</td>
<td>3.0</td>
<td>.103</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.00 (10)</td>
<td>2.57 (7)</td>
<td>2.71</td>
<td>6.0</td>
<td>.035*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training

- will help generate management support
- will help produce more effective training
- would be feasible in their organization

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>ULM: Yes</th>
<th>ULM: No</th>
<th>TD: Yes</th>
<th>TD: No</th>
<th>t-value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3.56 (10)</td>
<td>3.50 (8)</td>
<td>0.11</td>
<td>15.0</td>
<td>.914</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.36 (15)</td>
<td>2.85 (13)</td>
<td>1.19</td>
<td>25.0</td>
<td>.245</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.69 (17)</td>
<td>3.33 (6)</td>
<td>0.93</td>
<td>20.0</td>
<td>.366</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 20: Continued

<table>
<thead>
<tr>
<th>Evaluation Activity</th>
<th>Direction of Disagreement</th>
<th>t-Value</th>
<th>df</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collecting data (rate of sales, number and type of customer complaints, etc.) to</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>provide evidence of training impact</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>3.71 (7)</td>
<td>1.30</td>
<td>17.4</td>
<td>.211</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>3.82 (11)</td>
<td>1.44</td>
<td>28.8</td>
<td>.162</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>3.53 (15)</td>
<td>0.44</td>
<td>29.0</td>
<td>.665</td>
</tr>
<tr>
<td>Compare trainee skill or knowledge measurements before and after training to determine gains</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>3.38 (8)</td>
<td>1.91</td>
<td>23.0</td>
<td>.091</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>4.00 (4)</td>
<td>2.44</td>
<td>5.0</td>
<td>.058</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>4.00 (12)</td>
<td>4.18</td>
<td>11.0</td>
<td>.002*</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>2.67 (9)</td>
<td>1.21</td>
<td>19.4</td>
<td>.242</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>3.50 (4)</td>
<td>2.33</td>
<td>10.0</td>
<td>.042*</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>4.00 (6)</td>
<td>2.00</td>
<td>2.0</td>
<td>.184</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- will help generate management support</td>
<td>4.00 (10)</td>
<td>1.96</td>
<td>9.0</td>
<td>.081</td>
</tr>
<tr>
<td>- will help produce more effective training</td>
<td>3.33 (15)</td>
<td>-0.81</td>
<td>25.1</td>
<td>.427</td>
</tr>
<tr>
<td>- would be feasible in their organization</td>
<td>3.55 (11)</td>
<td>1.88</td>
<td>23.3</td>
<td>.073</td>
</tr>
</tbody>
</table>

*Numbers in parentheses indicate pairs of respondents.

*Probability of observing the reported difference in means for the sample when there is no difference in means in the population.
example, the mean for all situations in which the upper-level manager reported that specifically formulating training outcomes to reflect organizational need would be feasible and the training director reported it would not be feasible was 3.57 (on a 4-point scale, 1 being the most frequent and 4 being the least frequent). In this instance, when the training director said this activity would be feasible and the upper-level manager said it would not be feasible, the mean was 2.00. If a relationship existed between the direction of the disagreement and the frequency with which the evaluation was occurring, it would be reflected in the means. A t-test was used to determine if the differences for the groups could have been the result of a sampling error.

The population used was drawn from the matched responses. A matched response occurred when the researcher received a response from both the upper-level manager and the training director in the same organization. In 30 out of 33 instances, or 91% of the time, the mean score for evaluation activity showed a greater frequency when it was the training director rather than the upper-level manager who agreed with the value of, need for, or feasibility of the evaluation activity. A t-test run on the mean scores showed that those indicated by an asterisk had a probability of more than 95% of being true differences in the population.

The results of the study, as shown in Table 20, show no support for Hypothesis 4. The results indicate that, for all evaluation activities perceive to help generate management
support, it is the positive perception of the training director toward the activity that seems to be related to the frequency with which the activity occurs. In activities 2 and 4, the t-test results show the difference in the means is probably a real differences in the population and is not due to sampling error.

The results of the study, as shown in Table 20, do not support Hypothesis 5. In 9 of 11 cases the mean score for frequency indicated more activity when the training director showed positive perceptions of the activity than when the upper-level manager showed positive perceptions of the activity. In 3 of 11 cases, the t-test scores show that the difference in means is probably a real difference in the population as well as in the sample.

The results of the study show no support for Hypothesis 6. In 10 of 11 evaluation activities, the mean score for frequency indicated a greater level when the training director agreed that the activity was feasible and the upper-level manager disagreed than when the upper-level manager agreed and the training director disagreed. In 4 of 11 cases the difference in means, according to t-tests, was probably present in the population as well as in the sample.

In summary, there are no data to support the theory that when there is a disagreement between the upper-level manager and the training director regarding value, need, or feasibility of training-evaluation activities, the upper-level manager is the most
influential in determining whether or not the activity is actually carried out. The positive perception of the training director on value of, need for, or feasibility of an evaluation activity seems to have the greater influence on frequency of activity when there is disagreement.

A Comparison of the Frequency With Which Evaluation Activities Were Accomplished Between Respondents Who Were and Those Who Were Not Satisfied With the Quality of Their Training Program Evaluations

Training directors were asked if they were satisfied with the quality of their training program evaluations. The response choices were Yes; No; and N/A, we don't evaluate. Fifty-three training directors responded that they were satisfied with their training program evaluations, 48 said they were not satisfied with such evaluations, and 3 indicated they did not evaluate. On Part II of the survey, the training directors were asked to estimate the frequency with which they accomplished 11 training-evaluation activities. They were asked if they used a particular evaluation activity usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), or hardly ever (0-24% of the time). Responses were then coded, with 1 being the most frequent (usually) and 4 being the least frequent (hardly ever). The means were calculated separately for each group: training directors who were satisfied with the quality of their training-program evaluations and those who were not satisfied. A t-test was run on the means to
determine the probability of observing the reported difference in means for the sample when there is no difference in means for the population. Numerical differences between the means would be expected, but because of the amount of variance, the differences might not be statistically significant.

Table 21 reports the means, the t-values, the degrees of freedom, and the probability that the means did not represent a real difference in the population. The differences in the means showed a greater frequency of training-evaluation activities for the training directors who were satisfied with their training evaluation. Six of the 11 means were significant to at least the .05 level. Four of those six were differences in means for evaluation activities involving the worth of training.

A Comparison of the Frequency With Which Evaluation Activities Were Accomplished Between Respondents Who Were and Those Who Were Not Required to Report Evaluation Results to Upper Management

Training directors were asked, "If you share training evaluation results with anyone in management above you, do they require it?" The response choices were Yes or No. Of the 100 training directors who responded to the question, 32 said they were required to share evaluation results, and 68 said they were not required to do so.
Table 21
A Comparison of Evaluation-Activity Means Between Training Directors Who Were Satisfied With the Quality of Training-Evaluation Activity and Training Directors Who Were Not Satisfied With the Quality of Training-Evaluation Activity (N = 101)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>1.28 1.54</td>
<td>-1.55</td>
<td>78.5</td>
<td>.125</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>1.92 2.58</td>
<td>-3.07</td>
<td>98.0</td>
<td>.003*</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>2.38 3.02</td>
<td>-2.69</td>
<td>98.0</td>
<td>.008*</td>
</tr>
<tr>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their skills/knowledge on the job</td>
<td>2.50 3.25</td>
<td>-3.20</td>
<td>98.3</td>
<td>.002*</td>
</tr>
<tr>
<td>Monitoring training during the training activity to make course corrections</td>
<td>2.28 2.46</td>
<td>-0.80</td>
<td>97.8</td>
<td>.428</td>
</tr>
<tr>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>2.67 3.02</td>
<td>-1.49</td>
<td>98.0</td>
<td>.138</td>
</tr>
<tr>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>2.90 3.42</td>
<td>-2.49</td>
<td>97.0</td>
<td>.014*</td>
</tr>
<tr>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>3.08 3.56</td>
<td>-2.61</td>
<td>93.6</td>
<td>.010*</td>
</tr>
<tr>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>2.71 3.15</td>
<td>-2.04</td>
<td>98.0</td>
<td>.043*</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td>2.38 2.44</td>
<td>-0.26</td>
<td>99.0</td>
<td>.797</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>3.23 3.52</td>
<td>-1.65</td>
<td>99.0</td>
<td>.102</td>
</tr>
</tbody>
</table>

*Probability of observing the reported difference in means for the sample when there is no difference in means for the population.
On Part II of the survey training directors were asked to estimate the frequency with which they accomplished 11 training-evaluation activities. The response choices were usually (75-100% of the time), often (50-74% of the time), sometimes (25-49% of the time), and hardly ever (0-24% of the time). Responses were coded, with 1 being the most frequent (usually) and 4 being the least frequent (hardly ever). The mean performance frequency was then calculated for both training directors who were required to report evaluation results to upper management and those who were not required to do so. A t-test was run on the means to determine the probability of observing the reported difference in means for the sample when there is no difference in means for the population.

Table 22 reports the means, t-values, degrees of freedom, and probability for each of the 11 training-evaluation activities. The differences in means showed that there was more frequent training-evaluation activity for training directors who were required to report training-evaluation results than for those who were not required to report such results. The first and the second activities showed a significance of at least .05. This means there was at least a 95% chance that the difference in means was real in the population as well as in the sample.
Table 22
A Comparison of Evaluation-Activity Frequency Between Respondents Required to Report Training-Evaluation Results to Upper Management and Those Not Required to Report Training-Evaluation Results to Upper Management (n = 100)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td>1.13, 1.53</td>
<td>-2.97</td>
<td>97.9</td>
<td>0.004*</td>
</tr>
<tr>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td>1.94, 2.38</td>
<td>-1.99</td>
<td>69.9</td>
<td>0.050*</td>
</tr>
<tr>
<td>Taking trainee skill or knowledge measurements during training programs</td>
<td>2.50, 2.74</td>
<td>-0.93</td>
<td>97.0</td>
<td>0.354</td>
</tr>
<tr>
<td>Following up after trainees have gone back to their jobs (3-12 mos.) to see if they are able to use their skills/knowledge on the job</td>
<td>2.63, 2.99</td>
<td>-1.34</td>
<td>49.9</td>
<td>0.187</td>
</tr>
<tr>
<td>Monitoring training during the training activity to make course corrections</td>
<td>2.09, 2.47</td>
<td>-1.62</td>
<td>98.0</td>
<td>0.109</td>
</tr>
<tr>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td>2.69, 2.90</td>
<td>-0.78</td>
<td>52.8</td>
<td>0.440</td>
</tr>
<tr>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td>3.00, 3.21</td>
<td>-0.91</td>
<td>96.0</td>
<td>0.365</td>
</tr>
<tr>
<td>Collecting data (rate of sales, etc.) to provide evidence of training impact</td>
<td>3.06, 3.41</td>
<td>-1.68</td>
<td>98.0</td>
<td>0.096</td>
</tr>
<tr>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td>2.78, 2.97</td>
<td>-0.81</td>
<td>97.0</td>
<td>0.418</td>
</tr>
<tr>
<td>Soliciting judgments about training plans from participants</td>
<td>2.09, 2.53</td>
<td>-1.77</td>
<td>98.0</td>
<td>0.080</td>
</tr>
<tr>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td>3.31, 3.47</td>
<td>-1.80</td>
<td>98.0</td>
<td>0.117</td>
</tr>
</tbody>
</table>

*Probability of observing the reported difference in means for the sample when there is no difference in means for the population.
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The purpose of this chapter is to summarize the research problem, methodology, procedures, and results of this study. On the basis of the results, general conclusions, recommendations, and implications are stated.

Summary

To train, educate, and develop employees, corporations alone are spending at least $30 billion annually. Fifteen million persons or one in every eight working Americans are involved in 17.6 million formal training programs each year. Just as they are making capital investments, corporations are investing in workers' knowledge, skills, and attitudes to create, as an end product, more productive employees. Investments in training, education, and development, like investments in other resources, are expected to produce a return of value to the organization. To demonstrate that value, training practitioners must evaluate both the merit and worth of their training programs.

The evaluation literature offers a variety of evaluation models that can be applied to the corporate setting. However, despite the relationship between evaluation activity and the design and delivery
of training programs, and despite the proliferation of literature and models, there is widespread agreement that the use of evaluation lags far behind the encouragement it has been given in the professional literature.

The primary purpose of this study was to examine the reasons for the discrepancy between the emphasis on training-program evaluation in the literature and actual practice in the field. This was accomplished by studying the perceptions of two stakeholders in the training process: upper-level managers and training directors. Specifically, these groups' perceptions were compared as to (a) the perceived value of each training-evaluation activity in generating management support, (b) the need for training-evaluation activities in helping produce more effective training, and (c) the feasibility of conducting training-evaluation activities in their organizations. In addition to comparing perceptions, the writer gathered data on the frequency with which 11 training-evaluation activities were accomplished. The results of this study will enable practitioners to gain insight into the reasons for the gap between the encouragement given to training-program evaluation in the professional literature and actual practice.

Six directional hypotheses were formulated. Three were intended to investigate whether there was a relationship between the frequency with which training-evaluation activities were carried out (dependent variable) and the congruence in the perceptions of upper-level managers and training directors regarding the value of
evaluation activities in generating management support, the need for evaluation activities in helping produce better training, and the feasibility of conducting training-evaluation activities in their organizations (independent variable).

Three additional hypotheses postulated that, in the event of disagreement about the value of evaluation activities in generating management support, the need for evaluation activities in helping produce better training, and the feasibility of conducting training-evaluation activities, it was the positive agreement of upper-level management (independent variable) that would be related to an actual increase in the frequency of training-evaluation activity (dependent variable).

The review of related literature assisted in forming a basis and impetus for the study. Evaluation models designed to measure both the merit and worth of training were highlighted, and current practices were investigated. Current literature indicated that most training evaluations measure the degree to which trainees and their supervisors accept the training program. Of evaluations that go beyond trainees’ or supervisors’ reactions, the majority evaluate trainees’ performance against predetermined criteria immediately after training.

Finally, the need for, value of, and feasibility of training evaluation were discussed. The need and value issues are related to the question, "Why evaluate?" There must be some payoff with regard
to evaluation outcomes. The third issue, feasibility, is related to
the question, "Why not evaluate?" The feasibility question deals
with deterrents and how possible it is to achieve conditions that
allow for program evaluation.

A sample of 323 nonmanufacturing business organizations with
1,000 or more employees was selected from Dunn and Bradstreet's
(1986) Million Dollar Directory, using a systematic sampling
procedure. The writer identified two respondents from each
organization: a training director or someone who managed the
organization's training function and the person to whom that
training director reported, referred to as an upper-level manager.
All 323 organizations were contacted; 165 provided names of
potential respondents. Of those that chose to provide names, 48
had policies against participating in research, 27 had gone out of
business or were in the process of reorganizing, and 17 were holding
companies. Sixty-six organizations had no identifiable training
director or department. Of the 165 organizations that provided
names of potential respondents, 83 were represented by surveys from
both a training director and an upper-level manager, 21
organizations were represented by a training director only, and 16
were represented by an upper-level manager only. All returned
surveys were usable.

Each participant was asked to supply demographic information
regarding the organization he/she represented. Upper-level managers
were asked for the name of the organization, principal product or
service offered, type of organization, and approximate number of employees in the organization. In addition, upper-level managers were asked how important training was to the success of the organization. Training directors were asked questions to determine the scope of training-evaluation activities. Six questions concerned the kind of training evaluated, satisfaction with training evaluation, motivation for training evaluation, who conducted training evaluation, and with whom training-evaluation results were shared.

In the second part of each questionnaire, both groups were given a list of 11 training-evaluation activities. These activities mirrored Brinkerhoff's (1987) six-stage evaluation model. For each of the 11 items, respondents were asked to judge whether or not the activity would (a) produce information that would help generate management support for training, (b) provide information needed to help produce more effective training, and (c) be feasible in their organization.

The training directors were also asked to judge, on a 4-point scale, whether they usually, often, sometimes, or hardly ever used that activity for training evaluation.

The demographic data and information about the scope of the training programs and training-evaluation activities were reported in frequencies and percentages. Rank ordering was given when appropriate.
A chi-square test, Cramer's $V$, and lambda were used to measure the association between positive agreement, negative agreement, and disagreement on the part of training directors and upper-level managers regarding the value of, need for, and feasibility of training-evaluation activities (independent variable) and the degree to which training-evaluation activities were generally carried out (dependent variable). In addition, $t$-tests were performed to discover whether there were differences in frequency of evaluation activities between (a) training directors who were satisfied with the quality of their training-evaluation activities and those who were not satisfied, and (b) training directors who were required to report evaluation-activity results to upper-level management and those who were not required to do so.

The limitations mentioned in Chapter III regarding perceptual data were minimized. There are at least two reasons to believe the data collected reflect reality. First, the variability in types of responses indicated that subjects were discriminating among the activities. Second, the data collected in this study, particularly regarding the frequency of evaluation activities, approximated the findings of Brandenburg (1982), the Bureau of National Affairs (1985), Clegg (1978), and Sullivan (1970).

Conclusions

The following conclusions based on the data collected in this study are recognized as being characteristic of nonmanufacturing
organizations having an identifiable training unit and more than 1,000 employees.

Conclusions Related to Hypotheses 1 Through 3

None of the first three hypotheses was decisively supported by statistical significance as reflected in the data. The first three hypotheses stated that it is the positive congruence in perceptions between the upper-level manager and the training director concerning the value of, need for, and feasibility of training-evaluation activities that is related to the frequency with which training-evaluation activities are conducted. Several conclusions were drawn about why the writer was unable to support the first three directional hypotheses.

Conclusion 1: Training remains a corporate exercise that is taken on faith; little or no demand has been made to evaluate it in a rigorous fashion. Training is assumed to enhance the purpose and goals of the organization.

It is possible, as Wallace and Twitchell (1953) believed, that there is a lack of training evaluation because management does not want to waste time evaluating something it is already convinced is good. "Training programs are uniformly excellent by expert opinion and proclamation" (Wallace & Twitchell, 1953, p. 25). Of the upper-level managers responding in this study, 91.9% reported that training was very to extremely important to their organization's success. Yet only 32% of the training directors reported they were
required to share evaluation results with anyone in management above them.

One might conclude, then, that although Table 15 showed a great deal of agreement between training directors and upper-level managers concerning the value of, need for, and feasibility of training-evaluation activities, those activities were not required by upper-level managers. Tables 11 and 12 reinforced that conclusion by showing that training directors did not consider requirement by a higher authority to be a significant factor in motivating them to accomplish training evaluation.

**Conclusion 2:** Management does not see training as a management tool and is not asking what training has contributed to the achievement of organizational goals.

Although upper-level managers and training directors agreed that the five evaluation activities measuring worth were most likely to help generate management support, 30% or less of the training directors reported generally carrying out these measures of worth. Only 18.1% of the training directors generally gathered data to provide evidence of the effect of training on the organization. Until training must be submitted to cost/benefit analysis, management will continue to use or discard a tool of unknown worth.

**Conclusion 3:** Although none of the hypotheses was supported, the indication was that the strongest relationship existed between the perceived feasibility of the training-evaluation activity and the frequency with which training evaluation occurred.

The four items seen as most feasible (Table 15, column 2) were also the four items most frequently carried out (Table 15, column
Similarly, the middle four items in the feasibility ranking (Table 15, column 2) were also the middle four items in terms of frequency with which they were carried out. The three items seen by all respondents as least feasible (Table 15, column 1) were also those that were least frequently carried out.

Agreement on the ability of an activity to produce more effective training or increased management support seemed to be unrelated to whether or not the activity was carried out. As shown in Table 15, column 3, there was a high level of agreement (94.6%) that following up after trainees had gone back to their jobs (3 to 12 months after training) to see if they were able to use their skill/knowledge on the job would help produce more effective training. And yet, only 30.1% of all training directors actually accomplished this evaluation activity (Table 15, column 1).

As seen in Table 15, column 4, 84.2% of the upper-level managers and training directors agreed that making judgments about anticipated costs of training as compared to anticipated effectiveness of training would help generate management support for training. Only 24.1% of all training directors actually accomplished this task (Table 15, column 1).

Conclusions Related to Hypotheses 4 Through 6

None of the last three hypotheses was conclusively supported by statistical significance as reflected in the data. Hypotheses 4
through 6 stated that where there is disagreement in perceptions between the upper-level manager and the training director concerning the value of, need for, or feasibility of accomplishing training-evaluation activities, it is the positive perceptions of the upper-level manager that are related to the frequency with which the activity is carried out.

**Conclusion 4:** The positive perception of the training director concerning the value of, need for, or feasibility of an evaluation activity seemed to have greater influence on the frequency with which that activity was carried out than did the perception of the upper-level manager when a disagreement between the upper-level manager and training director existed.

In 30 out of 33 instances or 91% of the time, the mean score for frequency of the evaluation activity was higher when the training director rather than the upper-level manager agreed with the value of, need for, or feasibility of that activity (Table 20). A t-test run on the scores showed that 10 of the 30 instances were statistically significant at the .05 level. This finding provides further evidence that it is not upper management that makes demands on training, nor does it ask what training has contributed to the achievement of organizational goals. Rather, the training director chooses the evaluation activities and determines when they will apply. The needs of the organization, as filtered through upper management, are not pervasive in the training environment.

**Conclusion 5:** Although none of the hypotheses was supported, the indication was that for those organizations in which there was positive agreement concerning evaluation activities that help generate management support, help produce more effective training, or are feasible in the organization, there was also greater frequency of training-evaluation activities.
Of 33 agreement/disagreement opportunities for training directors and upper-level managers (Table 16), in 82% (27 of 33) of the cases in which there was positive agreement, the mean score for frequency of the training-evaluation activity was lower (indicating more activity on the 4-point scale) than in cases of disagreement or negative agreement. In this sample, the congruent perceptions of the training director and the person to whom he/she reported seemed to be related to an increased amount of training-evaluation activity.

**Conclusions Unrelated to the Hypotheses**

In addition to conclusions related to the hypotheses, several other conclusions were drawn.

**Conclusion 6:** Training directors' lack of satisfaction with training evaluation may have been associated with their not having completed the more complex evaluation of the worth of training.

When the frequency means for 11 evaluation activities of training directors who reported they were satisfied with their training-evaluation efforts were compared with the frequency means for 11 evaluation activities of training directors who were not satisfied with their training-evaluation efforts, the satisfied training directors consistently showed more frequent activity. The differences were significant for six of the activities. Four of the six were differences in activities measuring worth (Table 21).
Only three measurements of merit were generally accomplished by the dissatisfied training directors. These were: end-of-training evaluation concerning trainee satisfaction with the training experience, monitoring training during the training activity to make course corrections, and soliciting judgments about training plans from participants. The satisfied training directors also generally accomplished those three merit-evaluation activities but added one additional merit-activity (taking trainee skill knowledge measurements during training) and two activities measuring worth (specifically formulating training outcomes to reflect organizational need and following up after trainees had gone back to their jobs to see if they were able to use their skills/knowledge on the job). Perhaps training directors understood what constitutes good evaluation and recognized when their evaluation activities were inadequate.

Conclusion 7: Most evaluation efforts were focused on the immediate aspects of training and did not reflect the context of the total organization.

Most evaluation efforts dealt with good training as the result when they should have been dealing with the effectiveness or profitability of the organization. When ranking factors motivating them to evaluate training, training directors indicated that determining the effectiveness of training and improving training designs were the two strongest motivators (Tables 11 and 12). Brandenburg (1982) found that professional trainers thought the primary function of training evaluation was to improve the training
programs. These perceptions clearly indicated that the purpose of evaluation, as perceived by professional trainers, was to produce better training. Brandenburg concluded from his study of professional training groups that evaluation techniques used most often were performed while participants were on-site in training and that these techniques were efficient, short, and not labor intensive. More complex techniques, requiring more time or personnel, were given lower priority and were used less often.

Further evidence to support this conclusion was found in the Bureau of National Affairs (BNA, 1985) report on training-evaluation activity. In that study, 7 out of 10 trainers conducted some kind of training evaluation; the most popular type of evaluation was soliciting participants' reactions immediately after training. In the present study, 85.5% of the respondents usually/often used evaluations of participants' reactions (Table 15, column 1). Almost all (96 to 99%) of the BNA participants who evaluated used participant reactions. Following up to measure changes in job performance was reported to be done usually/often by 30.1% of this study's participants. In the BNA study, slightly more than half of the participants measured changes in job performance. Few respondents in the present study (18.1%) and in the BNA study (25 to 27%) evaluated the effects of training on company goals.

The congruence of the findings of this study with those of various other investigations done over several years suggests that
the present findings have validity and can be generalized to different sites.

**Conclusion 8:** The frequency with which training-evaluation activities were reported to have been carried out lagged far behind the percentage of respondents who agreed that these activities were feasible in their organizations.

Eighty-one percent of all respondents said they thought that following up after trainees had gone back to their jobs to see if they were able to use their skill/knowledge on the job was a feasible activity in their organizations (Table 15, column 2). Yet only 30.7% of the training directors reported generally accomplishing this activity (Table 15, column 1). Similarly, 54.2% of the respondents saw collecting data to provide evidence of the effect of training as feasible in their organizations (Table 15, column 2), whereas only 18.1% of the training directors reported generally accomplishing this activity (Table 15, column 1).

**Conclusion 9:** Lack of time is still the most frequently cited impediment to the quality of training evaluation.

In Sullivan's 1970 study and Clegg's 1978 study, "lack of time" was reported by chief training officers as the major constraint to in-house evaluation efforts. "Lack of staff," second in this study, was not a choice in the Sullivan or Clegg studies. "Lack of adequate evaluation methods" was the second most frequently cited constraint in the Clegg and Sullivan investigations; in this study it was cited as the third most frequent impediment to quality training evaluation.
Conclusion 10: There may be reason to believe that when the training director is required to report training-evaluation results, more training-evaluation activity takes place.

A comparison of the frequency means for evaluation activities (Table 22) showed that there was more frequent training-evaluation activity for training directors who were required to report training-evaluation results than for those who were not required to report such results. For the first and second activities, the mean differences were significant at the .05 level.

Evaluation results were most often shared with others in the training department (87.1%), the director's manager (87.1%), and the person conducting the training (82.2%). However, most of the directors (68%) said they were not required to share evaluation results with upper management.

Recommendations

A number of important issues emerged as a result of this study. The following comments and recommendations are directed to training directors, training managers, and upper-level managers.

Recommendation 1: Training directors or those who directly manage the training function should take the lead in establishing the worth of their efforts.

If it is primarily the training director who determines which training-evaluation activities will take place, as the findings of this study suggest (Table 20), then it is the training director who has the opportunity to take the lead in establishing the worth of training.
Evaluation, as a tool rather than an end product, is a means to understanding. The role evaluation presently plays was reflected in the categories of most frequently collected data (Table 15). Those data still reflected the notion that good training is the desired end product. Evaluating training for the purpose of producing better training is a narrow and dangerous focus and could easily lead to doing well what is not worth doing at all.

Improvement in human performance to achieve the goals of the organization is the actual end product of training. Training directors must relate their activities to organizational achievement or risk being discarded as a tool of unknown worth.

Recommendation 2: Training directors might investigate a more imaginative approach to solving the problem of "too little time."

Rather than trying to conduct training-evaluation activities by themselves, as this study indicated is usually the case, training directors might rather use supervisors of trainees or other interested parties to design evaluation measures and gather useful data. Serious consideration should be given to "training audits" conducted by someone outside the training department. The "audit" should take into account all six stages of evaluation outlined by Brinkerhoff (1987).

Recommendation 3: Upper-level managers should demand accountability in terms of the contribution of training to organizational goals.

Because more than 50% (and, in 8 of 11 cases, more than 70%) of all training directors and upper-level managers thought that any of
the evaluation activities were feasible in their organizations (Table 15), there is reason to believe that if upper-level managers demanded accountability through training evaluation, such evaluation would take place more frequently.

Recommendation 4: An organization whose training department has been disbanded should be examined to determine if the department was terminated because it failed to establish its worth.

In this study, 91.9% of the upper-level managers reported that training was very to extremely important to their organization's success. Yet only 32% of the training directors were required to provide any proof of success. Less than 19% generally made any attempt to provide evidence of the contribution of training to the organization's success (Table 15). Investigating why training departments were disbanded may provide training directors with incentive to take the lead in demonstrating their worth.

Recommendation 5: Training directors who accomplish training-evaluation activities related to worth should be part of a separate study to determine what sustains and reinforces their activities.

If it is not necessarily the upper-level manager who is providing the motivation for evaluating the worth of training, as this study suggests, then further research should be conducted to determine what in the training director's environment is providing the reinforcement for the time and resource investment it takes to produce training evaluation demonstrating the worth of training.
Recommendation 6: Different techniques should be used to investigate the role of training evaluation in organizations.

To further validate the results of this study, it is recommended that other methods of investigation be used to determine the role of training evaluation in the business environment. An in-depth study of an organization whose training-evaluation activities measure the worth and merit of training and an in-depth study of an organization whose training-evaluation activities measure only the merit of training would be an alternative method of investigation.

Implications

This final section on implications reflects beliefs held by the writer regarding training and development. The beliefs expressed are not intended to relate directly to the data or hypotheses in the study. Rather, they reflect the writer's view of training and development from a perspective arrived at as a result of completing this study and having spent many years in education and training.

Given: Training can be effective in assisting managers to achieve business goals. Question: Is it? Answer: Who knows? Admittedly, the above sequence is simplified to highlight the problem. Let us dramatize the problem by considering the following series of events.

1. A line manager is frustrated by a condition in the workplace that she/he is convinced will be improved by training.
2. The training director responds by purchasing, developing, or applying an existing XYZ program.

3. Trainees are scheduled for the XYZ program.

4. At the end of the XYZ training program trainees are asked, "How did you like it?"

This all-too-typical series of events and others similar to it are flawed because of the lack of needs specification at the outset and results assessment at the conclusion.

Training is needs driven, which means that a critical part of determining worth begins with establishing intended organizational benefits. The questions to begin with are: Is the problem worth addressing? Can training effectively address the problem? Failing to evaluate with these questions at the beginning makes it unlikely that answers to other questions of worth can be pursued 3 to 6 months after the training intervention.

As demonstrated in this study, training directors who were dissatisfied with their training evaluations accomplished only a few merit-evaluation activities. This study also indicated that training directors are in control of whether or not evaluation activities will include measures of worth. If they are dissatisfied and in control, why are they not gathering data to provide evidence of worth?

Those who work in business understand that management support is critical to survival in the organization. In this study, training directors clearly indicated they understood that measures of
worth, above measures of merit, help generate management support. Training directors also indicated that many measures of worth were feasible in their organizational settings. If management support is critical and measures of worth are known to generate management support and they are feasible, why are measures of worth not being provided to management?

The possible answer may be contained in yet another workplace scenario. This typical scenario consists of the busy training director doing what the company asks, accomplishing many activities, concentrating on providing good training programs, feeling successful, and not being required by his/her nontraining manager to prove the added value that results from training. Continuing to play out these scenarios and others like them most likely will result in doing well what may not be worth doing at all.

The results of this study showed that these scenarios are not being played out in ignorance. Training professionals know what needs to be done. However, a paradox exists in that their nontraining managers do not require them to produce evidence of the worth of training, whereas most important organizational functions are regularly required to evaluate bottom-line contributions. Perhaps we are waiting for our "scripts" to be rewritten.

As professionals, if we are waiting for our "scripts" to be written by upper management so that it parallels other important business functions in the organization, we are being naive and...
irrational and are missing an opportunity for self-determination. We have an opportunity to rewrite our own "scripts" and develop systems to supply upper management with data about needs, programs, plans, operations, and results. With such systems, training will assume a more proactive role and be recognized as a force that positively affects the outcome of the organization. The sequence we began with could then become: **Given:** Training can be effective in assisting management to achieve business goals. **Question:** Is it? **Answer:** Yes, and here is how we know....
APPENDIX A

SURVEY INSTRUMENTS
Part I includes questions about your organization and about training in general. They are intended to give us information so that we can group our respondents. Please read each item and respond accordingly.

1. Name of your organization _____________________________________________________________

2. Principal product/service offered _______________________________________________________

3. Type of organization (check one)
   - Merchandising or Retail Trade (i.e. department stores, hotels, restaurants, distributors, etc.)
   - Service Organization (i.e. transportation, insurance, financial institution, stock broker, newspaper, health care, etc.)
   - Utilities (i.e. power, water, telephone, etc.)
   - Entertainment (i.e. golf course, amusement park, theatre, etc.)
   - Other. Please specify ________________________________________________________________

4. Approximate number of employees in your organization ________________________________

5a. Is there at least one unit in your organization whose major responsibility is training?
   □ YES     □ NO

b. If yes, briefly describe the reporting relationship of the training director, manager, etc., through the CEO or President. For example: the Training Director reports to / the Vice President of Human Resources reports to / the Senior Vice President of Operations reports to / the President

   If possible enclose a chart showing the reporting relationship of the training unit or department.

6. How do you view the importance of training to your organization’s success?
   □ Extremely important
   □ Very important
   □ Neither important nor unimportant
   □ Not very important
   □ Not at all important
Part II

On the next page are activities related to training evaluation. Some activities related to the delivery and evaluation of training programs are more suitable to your situation than others. Please think about each activity and decide if it is an activity that:

- would produce information that generates management support for training.

- would provide information to help produce more effective training.

- would be feasible in your organization and could be successfully carried out.

Please apply these definitions as you respond.

FOR EXAMPLE:

If you felt the activity below,

— would not necessarily help generate management support for training,

— but would help produce better training programs,

— and as you consider the activity, it could be successfully carried out in your organization,

your response would look like this:

<table>
<thead>
<tr>
<th>would help generate management support</th>
<th>would help produce more effective training</th>
<th>would be feasible in this organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Analyzing the performance of high-producers and comparing it with low-producers to determine the content of training: ✓ ✓ ✓
Check the box which best represents your view. Remember this is your opinion only and does not reflect whether or not your organization actually performs these activities. Be sure to check a box in each of the three categories for all 11 items.

<table>
<thead>
<tr>
<th>Number</th>
<th>Activity</th>
<th>Would help generate management support</th>
<th>Would help produce more effective training</th>
<th>Would be feasible in this organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>End-of-training evaluation concerning trainee satisfaction with the training experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Specifically formulating training outcomes to reflect organizational need</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Taking trainee skill or knowledge measurements DURING training programs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Following up after trainees have gone back to their jobs (3-12 months after training) to see if they are able to use their skills/knowledge on the job</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Monitoring training during the training activity to make course corrections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Taking skill or knowledge measurements after training to determine trainee achievement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Collecting data (rate of sales, number and type of customer complaints, employee turnover, 22% of quoted, etc.) to provide evidence of training impact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Comparing trainee skill or knowledge measurements before and after training to determine gains</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Soliciting judgments about training plans from participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Gathering data to compare training costs with the organizational impact of training</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

May I list your company as having participated in this study?

[ ] YES   [ ] NO

Are you interested in receiving a copy of the results of this study?

[ ] YES   [ ] NO

THANK YOU for your time and responses. Please return the completed inventory (in the enclosed, stamped, self-addressed envelope) to:

Sarah Cutak
1122 Ottawa Ave.
Grand Haven, Michigan 49417

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
**B Part I**

The following are questions about training in general. They are intended to give us information so that we can appropriately group our respondents. Please read each item carefully and respond accordingly.

(1) Approximately how many of your employees participated in some form of training provided by your training unit during the past year?
- [ ] 0
- [ ] 1-100
- [ ] 100-300
- [ ] 300-500
- [ ] 600-800
- [ ] 900 or more

(2) Of the training for which you are responsible, what percentage is:
- [ ] % delivered by you or your training staff (in-house training)
- [ ] % provided by persons from outside of your organization (external consultants)
- [ ] % informal, on-the-job training
- [ ] N/A. Please explain

The following six questions ask about the evaluation of training programs. Evaluation might include activities such as: trainee reaction, post-training testing, supervisor or manager judgments, instructor reactions, anecdotal records, etc.

(3) Do you evaluate:
- [ ] YES, ALWAYS
- [ ] YES, SOMETIMES
- [ ] NO, NEVER

<table>
<thead>
<tr>
<th>Training conducted by you or your staff?</th>
<th>YES</th>
<th>SOMETIMES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training conducted by external consultants?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informal, on-the-job training?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(4) a. Are you satisfied with the quality of the training program evaluations that you do?
- [ ] YES
- [ ] NO
- [ ] N/A we don’t evaluate (go to Part II, page three of the survey)

b. If you are not satisfied with the quality of your training program evaluation, please indicate those items that may be impediments to evaluation (check all that apply)

<table>
<thead>
<tr>
<th>Lack of budget</th>
<th>Not required by a higher authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of time</td>
<td>Lack of staff</td>
</tr>
<tr>
<td>Lack of expertise</td>
<td>Not considered important by the training manager</td>
</tr>
<tr>
<td>Lack of adequate evaluation methods</td>
<td>Other. Please specify...</td>
</tr>
</tbody>
</table>
(5) If your answer to question three was YES, ALWAYS or YES, SOMETIMES for any category of training programs, think of those programs that are evaluated and why they are evaluated. Rank the following in order of their importance in motivating you to evaluate training. Rank ALL of the items using ONE (1) as the HIGHEST, most motivating factor. No ties please.

a. ___ required by a higher authority
b. ___ used to justify training programs
c. ___ used to determine the effectiveness of training
d. ___ used to improve the training design for future programs
e. ___ used to measure the application of training back on the job
f. ___ used to provide feedback for trainees on their progress in knowledge and/or skills
g. ___ used to measure progress toward the company's objectives/profit
h. ___ used to provide evidence of training's cost effectiveness
i. ___ other. Please specify ________________________________

(6) Who GENERALLY conducts the evaluation of training programs? Check all that apply.

<table>
<thead>
<tr>
<th>the training staff</th>
<th>training conducted by you or your staff</th>
<th>training conducted by external consultants</th>
<th>informal, on-the-job training</th>
</tr>
</thead>
<tbody>
<tr>
<td>an ad hoc committee</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in-house evaluation experts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>an outside consultant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>supervisors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>others (please specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(7) With whom do you share evaluation results? Check all that apply.

a. ___ others in the training department
b. ___ the person to whom I report
c. ___ trainees
d. ___ supervisors or managers of the trainees
e. ___ supervisors requesting the training
f. ___ the person conducting the training
g. ___ others. Please specify ________________________________

(8) If you share training evaluation results with anyone in management above you, do they require it?

☐ YES  ☐ NO
Part II

On the next page are activities related to training evaluation. Some activities related to the delivery and evaluation of training programs are more suitable to your situation than others. Please think about each activity and decide if it is an activity that:

- would produce information that generates management support for training,
- would provide information to help produce more effective training,
- would be feasible in your organization and could be successfully carried out, and
- you usually (75-100% of time), often (50-74% of time), sometimes (25-49% of time) OR hardly ever (0-24% of time) do it that way now.

Please apply these definitions as you respond.

FOR EXAMPLE:

If you felt the activity below,
- would not necessarily help generate support for training,
- but would help produce better training programs,
- and as you consider the activity, it would be feasible for you to do it now,
- but, in reality you do not do it now,

your response would look like this:

<table>
<thead>
<tr>
<th>Analysing the performance of high-producers and comparing it with low-producers to determine the context of training.</th>
<th>would help generate management support</th>
<th>would help produce more effective training</th>
<th>would be feasible in this organization</th>
<th>In the way we do it</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Check the box which best represents your view and/or situation. Think only of the training for which your unit is responsible. Be sure to check a box in EACH of the four categories for all 11 items.

<table>
<thead>
<tr>
<th>Would help generate management support</th>
<th>Would help produce more effective training</th>
<th>Would be feasible in this organization</th>
<th>In the way we do it</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

1. End-of-training evaluation concerning trainee satisfaction with the training experience.
2. Specifically formulating training outcomes to reflect organizational need.
3. Taking trainee skill or knowledge measurements DURING training programs.
4. Following up after trainees have gone back to their jobs (3-12 months after training) to see if they are able to use their skills/knowledge on the job.
5. Monitoring training during the training activity to make course corrections.
6. Taking skill or knowledge measurements after training to determine trainee achievement.
7. Making judgments about the anticipated costs of training as compared to the anticipated effectiveness of training.
8. Collecting data (rate of sales, number and type of customer complaints, employee turnover, rate of output, etc.) to provide evidence of training impact.
9. Comparing trainee skill or knowledge measurements before and after training to determine gains.
10. Soliciting judgments about training plans from participants.
11. Gathering data to compare training costs with the organizational impact of training.

Are you interested in receiving a copy of the results of this study?  

YES  NO

THANK YOU for your time and responses. Please return the completed inventory (in the enclosed, stamped, self-addressed envelope) to:

Sarah Gutek  
1132 Ottawa Ave.  
Grand Haven, Michigan 49417
APPENDIX B

LETTER AND POSTCARD SENT TO COMPANY PRESIDENTS
In today's competitive business environment, most companies are looking for ways to increase productivity. One of the ways of accomplishing increased productivity is through more focused internal training programs. As a doctoral student at Western Michigan University, I will be conducting a study which will identify differences between upper-level managers and training managers regarding the degree to which each feels training evaluation activities are feasible, needed, and valued. In order to begin this study I need your help. Would you be willing to participate?

Enclosed you will find a stamped, addressed post card on which you may provide:

- the name of the person in your company principally in charge of training (training manager, coordinator, specialist, etc.)
- the name of the person to whom the above named training manager, coordinator, specialist, etc. reports.

In January of 1987, I will send a short questionnaire to each of the persons listed on the card. The data produced by this study will in no way be tied to individual names or organizations. If you would like a copy of the survey results, check the appropriate space on the post card.

Thank you in advance for your willingness to cooperate.

Sincerely,

Sarah P. Gutek

Ken E. Dickie
Professor
Doctoral Advisor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
NAME OF YOUR COMPANY

In your company, who is the person principally in charge of training? (training manager, coordinator, specialist, etc.)

Name: ________________________
Title: ________________________

To whom does the above person report?

Name: ________________________
Title: ________________________

☐ We would like a copy of the survey results.
☐ We are not able to participate in the study, but would like a copy of the results.

THANK YOU!
APPENDIX C

COVER LETTER SENT WITH QUESTIONNAIRES
Earlier this month I called to request your company's participation in a study concerning training evaluation. Your name was given to me as the appropriate person to fill out one of the two questionnaires enclosed. Your manager should respond to Questionnaire A; you should respond to Questionnaire B.

The purpose of this study is to examine your perceptions and those of your manager regarding the need for, value of, and feasibility of selected evaluation activities. The results of this study will help us gain insight into the gap that exists between the encouragement given to training evaluation in professional literature and actual practice in business organizations. Your company will be one of approximately 150 companies participating nationwide.

The enclosed questionnaires ask for demographic data, information about existing evaluation practice, and your perceptions of selected evaluation activities. The value of this study rests with the comparison of each of your responses on PART II of the questionnaire; therefore, it is critical that you complete these parts independently. Pilot tests have shown that each questionnaire takes about 10 minutes to complete.

The data reported from this study will not be identified with individuals or organizations. The code number on the questionnaire will be used only for follow-up purposes. It is critical that both you and your manager return the questionnaires. A stamped, self-addressed envelope has been provided for each of you. If you have any questions, please call Sarah Gutek at (616) 942-3431.

If you would like to receive a copy of the results, please indicate that in the space provided at the end of the questionnaire. It would be most helpful if we could receive your responses by April 27.

Thank you for your time and cooperation.

Sincerely,

Sarah P. Gutek

Kenneth E. Dickie
Professor
Doctoral Advisor
Last December the president of your company indicated the company's willingness to participate in a study concerning training evaluation. Your name was given as the appropriate person to fill out the enclosed questionnaire.

The purpose of this study is to examine your perceptions and those of the person in charge of training in your organization regarding the need for, value of, and feasibility of selected evaluation activities. The results of this study will help us gain insight into the gap that exists between the encouragement given to training evaluation in professional literature and actual practice in business organizations. Your company will be one of approximately 150 companies participating nationwide.

The enclosed questionnaire asks for demographic data, information about existing evaluation practice, and your perceptions of selected evaluation activities. A similar questionnaire has been sent to the person in charge of your organization's training. The value of this study rests with the comparison of each of your responses on PART II of the questionnaire; therefore, it is critical that you complete this part independently. Pilot tests have shown that the questionnaire takes about 10 minutes to complete.

The data reported from this study will not be identified with individuals or organizations. The code number on the questionnaire will be used only for follow-up purposes. It is critical that both you and the person in charge of your organization's training return the questionnaires. A stamped, self-addressed envelope has been provided to each of you for that purpose. If you have any questions, please call Sarah Gutek at (616) 942-3431.

If you would like to receive a copy of the results, please indicate that in the space provided at the end of the questionnaire. It would be most helpful if we could receive your response by April 27.

Thank you for your time and cooperation.

Sincerely,

Sarah P. Gutek
Professor
Kenneth E. Dickie
Doctoral Advisor
Last December the president of your company indicated the company's willingness to participate in a study concerning training evaluation. Your name was given as the appropriate person to fill out the enclosed questionnaire.

The purpose of this study is to examine your perceptions and those of your manager regarding the need for, value of, and feasibility of selected evaluation activities. The results of this study will help us gain insight into the gap that exists between the encouragement given to training evaluation in professional literature and actual practice in business organizations. Your company will be one of approximately 150 companies participating nationwide.

The enclosed questionnaire asks for demographic data, information about existing evaluation practice, and your perceptions of selected evaluation activities. A similar questionnaire has been sent to your manager. The value of this study rests with the comparison of each of your responses on PART II of the questionnaire; therefore, it is critical that you complete this part independently. Pilot tests have shown that the questionnaire takes about 10 minutes to complete.

The data reported from this study will not be identified with individuals or organizations. The code number on the questionnaire will be used only for follow-up purposes. It is critical that both you and your manager return the questionnaires. A stamped, self-addressed envelope has been provided to each of you for that purpose. If you have any questions, please call Sarah Gutek at (616) 942-3431.

If you would like to receive a copy of the results, please indicate that in the space provided at the end of the questionnaire. It would be most helpful if we could receive your response by April 27.

Thank you for your time and cooperation.

Sincerely,

Sarah P. Gutek

Kenneth E. Dickie
Professor
Doctoral Advisor
Thank you for agreeing to participate in this study concerning training evaluation. Enclosed you will find two questionnaires. Your manager should respond to Questionnaire A; you should respond to Questionnaire B.

The purpose of this study is to examine your perceptions and those of your manager regarding the need for, value of, and feasibility of selected evaluation activities. The results of this study will help us gain insight into the gap that exists between the encouragement given to training evaluation in professional literature and actual practice in business organizations. Your company will be one of approximately 150 companies participating nationwide.

The enclosed questionnaires ask for demographic data, information about existing evaluation practice, and your perceptions of selected evaluation activities. The value of this study rests with the comparison of each of your responses on PART II of the questionnaire; therefore, it is critical that you complete these parts independently. Pilot tests have show that each questionnaire takes about 10 minutes to complete.

The data reported from this study will not be identified with individuals or organizations. The code number on the questionnaire will be used only for follow-up purposes. It is critical that both you and your manager return the questionnaires. A stamped, self-addressed envelope has been provided for each of you. If you have any questions, please call Sarah Gutek at (616) 942-3431.

If you would like to receive a copy of the results, please indicate that in the space provided at the end of the questionnaire. It would be most helpful if we could receive your responses by April 27.

Thank you for your time and cooperation.

Sincerely,

Sarah P. Gutek

Kenneth E. Dickie
Professor
Doctoral Advisor
APPENDIX D

SECOND COVER LETTER
As the saying goes, "There's good news and there's bad news". The good news is, I've received the questionnaire that I sent to your manager; the bad news is, I haven't received yours. Perhaps you've misplaced it or forgotten to fill it out. I am enclosing another copy of the questionnaire so you can share your perceptions and information regarding training evaluation.

Since the purpose of this study is to compare the perceptions of upper-level managers with those of training managers regarding the need for, value of, and feasibility of selected evaluation activities, I need to have both questionnaires returned to me in order to include your organization in the study. The enclosed questionnaire takes about 10 minutes to complete and your response is, of course, very important to the success of the study.

Your time and cooperation will be greatly appreciated.

Sincerely,

Sarah Gutek

Kenneth E. Dickie
Professor
Doctoral Advisor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
As the saying goes, "There's good news and there's bad news". The good news is, I've received the questionnaire that I sent to the person in charge of your organization's training; the bad news is, I haven't received yours. Perhaps you've misplaced it or forgotten to fill it out. I am enclosing another copy of the questionnaire so you can share your perceptions regarding training evaluation.

Since the purpose of this study is to compare the perceptions of upper-level managers with those of training managers regarding the need for, value of, and feasibility of selected evaluation activities, I need to have both questionnaires returned to me in order to include your organization in the study. The enclosed questionnaire takes about 10 minutes to complete and your response is, of course, very important to the success of the study.

Your time and cooperation will be greatly appreciated.

Sincerely,

Sarah Gutek

Kenneth E. Dickie
Professor
Doctoral Advisor
Just a friendly reminder that I need your participation in the research project regarding training evaluation.

Please complete and return the survey that was mailed to you on April 13th.

If you have already returned the survey, consider this card a thank you for your assistance. Sincerely, Sarah Gutek

(616) 942-3431
APPENDIX F

THIRD COVER LETTER
Perhaps you have misplaced or forgotten to complete the questionnaire you received from me about a month ago. As you may know, in a dissertation study of this type it is extremely important that I account for each of the randomly selected organizations, including yours. Consequently, I am enclosing another copy of the questionnaire and another self-addressed envelope. The questionnaire takes only 10 minutes to complete and your response is very important to the success of the study.

If you would like to receive a copy of the results of the study, please indicate that at the end of the questionnaire.

Your time and cooperation are essential and will be greatly appreciated.

Sincerely,

Sarah Gutek
(616) 942-3431

Kenneth E. Dickie
Professor
Doctoral Advisor
APPENDIX G

FOURTH COVER LETTER
I need your help! About a month ago I sent you two questionnaires: Questionnaire B was to be completed by the person in your organization who is in charge of employee training; Questionnaire A was to be completed by the person to whom that Training Director (Manager, Specialist, etc.) reports. Perhaps you have forgotten about or misplaced the questionnaires. As you may know, in a dissertation study of this type it is extremely important that I account for each of the randomly selected organizations, including yours. Consequently, I am enclosing a second set of questionnaires and self-addressed envelopes. Each questionnaire takes only 10 minutes to complete and these responses are very important to the success of the study.

You may receive a copy of the study's results by indicating so at the end of the questionnaire.

Your time and cooperation are essential and will be greatly appreciated.

Sincerely,

Sarah Gutek
(616) 942-3431

Kenneth E. Dickie
Professor
Doctoral Advisor
APPENDIX H

PARTICIPANTS WILLING TO BE CREDITED FOR RESPONDING
Abdow Big Boy Restaurants
Alfa Insurance Companies
Amalgamated Sugar Company
American Water Works, Incorporated
Amica Mutual Insurance Company
Anchor Savings Bank
Arizona Bank
Arthur J. Gallagher & Company
Bank of New York
Barnett Bank of Central Florida
Battelle Columbus Division
Belcan Corporation
Bonnie Be Lo Foods, Incorporated
Braniff, Incorporated
CB & T Bancshares, Incorporated
California Almond Growers Exchange
Centerre Bank
Chicago Northwestern Transportation
Citizens and Southern National Bank
Collins Foods International, Incorporated
Combustion Engineering
Community Mutual Insurance
Connecticut Bank and Trust Company
CUNA Mutual Insurance Group
Diamond M Company
Dime Savings Bank of New York
Dover Elevator Systems
DST Systems, Incorporated
Dunkin Donuts of America
Federal Reserve Bank of Cleveland
First Bank of Minneapolis/St. Paul
First National Bank of Chicago
Fluor Corporation
Foster Wheeler Corporation
General Accident Insurance Company
General Reinsurance Corporation
General Telephone Company of Ohio
Gilbane Building Company
Global Marine Drilling Company
Grand American Fare, Incorporated
Greyhound Food Management
Heller International
Hills Department Stores
Hilti Fastening Systems
IHOP Corporation
Integon Life Insurance
Leaseway Transportation Corporation
Lincoln Telecommunications Company
Magic Pan, Incorporated
Manpower, Incorporated
McDonnell Douglas Tymnet, Incorporated
Meijer, Incorporated
Michigan Consumers Power Company
Northeast Utilities
Nutri/Systems, Incorporated
Ohio Power Company
Pearle Health Services, Incorporated
Policy Management Systems Corporation
Progressive Casualty Insurance Company
Puget Sound Bank
Restaurant Associates Industries
Seaway Foodtown, Incorporated
Sea World of Florida
SEI Corporation
Southern California Gas Company
Southern National Bank of North Carolina
Southwestern Public Service Company
St. Johnsbury Trucking Company
Sundt Construction Company
Transamerica Insurance Company
Union Bank of California
United Illuminating Company
United Services Planning Association
Unitog Company
Whataburger, Incorporated
Wilson Industries, Incorporated
Wisconsin Public Service Corporation
Zions Co-Op Mercantile Institution
BIBLIOGRAPHY


