A STUDY OF THE IMPACT OF THE KAUFMAN ORGANIZATIONAL ELEMENTS MODEL FOR EVALUATION PRACTICES ON HUMAN RESOURCE DEVELOPMENT TRAINING PROGRAMS

By

Duane Gene Stevens

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology,
and Special Education

1988

ABSTRACT

A STUDY OF THE IMPACT OF THE KAUFMAN ORGANIZATIONAL ELEMENTS MODEL FOR EVALUATION PRACTICES ON HUMAN RESOURCE DEVELOPMENT TRAINING PROGRAMS

By

Duane Gene Stevens

This study had two major purposes:

1. To identify the percentage of each of four types of HRD training programs that are not evaluated in terms of organizational output, or outcome.

2. To determine why training evaluations are not done in terms of organizational output, or outcome.

The evaluation model underlying the study is the Kaufman Organizational Elements Model (OEM).

A structured telephone survey of a random sample selected from a population of HRD training managers and training evaluators was conducted. Participants were asked closed questions about the types of training programs offered by their departments, and the types of data collected for

training evaluation purposes. Participants were also asked open ended questions about the reasons why certain types of data were not collected. The estimated response rate was 47%. Ninety-nine people participated in the survey.

All research hypotheses for the closed questions were rejected on the basis of an analysis of proportions, indicating that there is a higher level of practice of the basic principles in the OEM Model than was hypothesized by the researcher. Alpha was set a priori at .05. Beta was at .005.

Reasons given for not performing training evaluations at the output level most often involved unperceived need for evaluation of training outputs, or that the idea of conducting evaluation of outputs had not even occurred to the person.

Reasons for not collecting data on training outcomes most often involved a perception that evaluation of outcomes is not needed, or that training outcomes are difficult to measure.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE PROBLEM

Significance of the Study

Training is undertaken by organizations to solve knowledge or skill deficiencies in the work force. Knowledge or skill deficiencies are one of the major sources of poor employee performance (Mager & Pipe, 1984). Poor employee performance results in poor productivity, shoddy quality control, expensive errors, client dissatisfaction, etc. The cost of knowledge and skill deficiencies is why over $100 billion dollars are spent annually by United States business and industry organizations (Corrigan, 1980).

Despite the major commitment to training in terms of expenditures, many training programs do not result in improved employee performance (Mager, & Pipe, 1984; Rummler, 1985). The failure of training programs to improve employee performance could be because:

- the training program was not really needed (i.e. there was no real performance problem to solve, or
- ineffective instructional strategies were employed, or
- employees were not allowed sufficient practice to acquire their new knowledge or skills, or

- rewards and punishments built into the work environment prevented proper performance, or
- employees were not provided with the proper support and tools to do the job correctly.

Evaluation of training programs must be carried out to determine if the programs are effective, and if not effective - why not. In other words, training evaluation should take place and should involve evaluating more than the degree to which learners enjoyed the training experience, or the degree to which learners actually mastered training objectives.

This study is significant in that it attempts to find out how much evaluation, and what kind of evaluation, is taking place within Human Resource Departments of organizations in the United States and Canada, and why certain kinds of evaluations are not being conducted. The author is not aware of any study to date that has identified the scope of evaluation practices in the United States and Canada, or reasons for not carrying out evaluations.

Over the last decade I have talked with many training managers and training evaluation specialists who believe that little training evaluation is actually conducted in organizations. In these conversations, the belief that training evaluations are not common practice is

given as a reason for lack of desire to carry out needed evaluations within their own organizations. Training managers or evaluation specialists may be mistaken in their beliefs relating to evaluation practices, or they may lack viable arguments to convince their management to invest funds into training evaluation.

This study should be of interest to HRD department managers because they need to know what constitutes state of the art evaluation practice in order to determine the types of evaluations that should be carried out in their departments. In addition, instructional evaluation specialists and training managers should find this study to be of importance because it identifies the degree to which certain types of evaluations are taking place for different types of training programs, and identifies the reasons why certain types of evaluations are not being done.

Purpose of the Study

This study had two major purposes:

1. To identify the degree to which human resource development (HRD) departments evaluate four different types of training programs in terms of the Kaufman Organizational Elements Model. Specifically, to identify the percentage of

each type of HRD training program that is not evaluated in terms of organizational product, output, or outcome.

2. To determine why training evaluations are not done with an emphasis on organizational product, output, or outcome.

The study data was gathered from the perspective of persons responsible for, or knowledgeable of, training evaluation practices within the HRD department in their organization. Gathering data from those most closely involved in HRD training evaluation was an important aspect of this study.

The Research Questions

The study was undertaken to answer two questions:

1. What is the degree of HRD department evaluation practices for each of the following types of training program:
   - Orientation
   - Wellness (health practices, exercise, diet, etc.)
   - Job Skill
   - Management Development

   in terms of organizational elements identified by Kaufman? Specifically, what percentage of training

programs within each of the above categories are evaluated in terms of products, outcomes, or outputs?

2. For any of the above types of training, if the training programs are not evaluated in terms of products, outcomes, or outputs, what rationale underlies the lack of such evaluation?

Hypotheses Tested

In order to investigate the first research question, the following hypotheses were formulated:

HO_1 Less than 5% of HRD orientation training programs are evaluated in terms of learner reaction to instruction.

HO_2 Less than 5% of HRD orientation training programs are evaluated in terms of learner mastery of objectives.

HO_3 Less than 1% of HRD orientation training programs are evaluated in terms of outputs.

HO_4 Less than 1% of HRD orientation training programs are evaluated in terms of outcomes.

HO_5 Less than 20% of HRD job skill training programs are evaluated in terms of learner reaction to instruction.

HO_6 Less than 10% of HRD job skill training programs are evaluated in terms of learner mastery of objectives.

HO_7 Less than 1% of HRD job skill training programs are evaluated in terms of outputs.

HO_8 Less than 1% of HRD job skill training programs are evaluated in terms of outcomes.

HO_9 Less than 20% of HRD management development training programs are evaluated in terms of learner reaction to instruction.

HO_{10} Less than 1% of HRD management development training programs are evaluated in terms of learner mastery of objectives.

HO_{11} Less than 1% of HRD management development training programs are evaluated in terms of outputs.

HO_{12} Less than 1% of HRD management development training programs are evaluated in terms of outcomes.

HO_{13} Less than 5% of HRD wellness training programs are evaluated in terms of learner reaction to instruction.

HO_{14} Less than 1% of HRD wellness training programs are evaluated in terms of learner mastery of objectives.

HO_{15} Less than 1% of HRD wellness training programs are evaluated in terms of outputs.

HO_{16} Less than 1% of HRD wellness training programs are evaluated in terms of outcomes.

The percentage estimates that form the basis of HO_1 through HO_{16} were not derived from a literature source or any given data base. I was not aware of any research in the areas of outputs or outcomes that could form the expected percentages for a statistical test. The percentage estimates contained in the research hypotheses were derived from my perception

of training evaluation practices. My perception is based upon information gained from a number of conversations with evaluation specialists over a course of several years. This is by no means a statistically reliable method of data gathering, but in the absence of hard data, contact with a number of professionals representing a cross section of organizations in the United States and Canada had to suffice.

Percentage estimates vary according to type of training program. I estimated that output and outcome evaluations would be highest for job skill training programs, and that output evaluations would be higher for management development training programs, but that in no case would a majority of training programs of a given type be evaluated in terms of products, outputs, or outcomes. The percentage estimates were shown to Roger Kaufman, and to Geary Rummler. Both evaluation theorists provided an opinion that the percentage estimates were not too low.

In order to investigate the second research question, the following hypotheses were formed:

HO_{17} The most frequently given reason for not evaluating any type of training program in terms of learner mastery is "Costs too much."

HO_{18} The most frequently given reason for not evaluating any type of training program in terms of output is "Would like to, but not enough budget."

HO_{19} The most frequently given reason for not evaluating any type of training program in terms of outcomes is "Can't be measured."

Again, I am not aware of any study to date that has identified the most frequent reasons for not performing evaluations for types of training programs, or for type of evaluation. Hypotheses HO_{17} through HO_{19} were based upon my intuition. Cost is often cited by training managers as a reason for not carrying out evaluations. The difference between HO_{17} and HO_{18} is that if evaluations of some sort are being carried out, but specifically not in terms of output, the training manager might desire to do so, but be limited by cost constraints. The difficulty of showing hard relationships between training and organizational outcomes (Kearsley, 1982), led to hypothesis HO_{19}.

Assumptions

The following were the basic theoretical or conceptual assumptions underlying this study:

1. That the persons to be contacted in this study would

possess the required knowledge to intelligently provide valid answers to the questions.

2. That respondents would answer truthfully, and would provide thoughtful answers to the questions.

3. That telephone interviews would result in thorough coverage of the questionnaire.

4. That personal telephone contact would result in a higher response rate than for mailed questionnaires.

5. That the respondents would understand the terminology used in the survey.

6. That identifying the degree to which evaluation is taking place within the context of the Kaufman Organizational Elements Model, or the reasons why not, is worthwhile.

Limitations

This study made no attempt to discover what methods were being used to carry out data collection. People merely reported that they gathered certain types of data as part of their evaluation process. The quality of the data gathering process used in evaluation efforts, or the degree to which

evaluation information effects training management decisions, was not determined in this study.

Due to limited funds available for the telephone survey, calling had to be cut off after 100 persons had been contacted. The limitation of 100 persons contacted resulted in sample sizes ranging from 22 to 63 for each type of training program. The small sample sizes resulted in sampling errors from .03 to .11.

Despite repeated attempts to reach prospective respondents, a large number could not be reached and had to be counted as non-responders. In addition a number of potential responders were reached but refused to participate in the study. The refusals resulted in an overall survey response rate of 47%, and contributed to the uncertainty of the representativeness of the data to the research population.

The telephone interviewers employed were not conversant in French or Spanish. As a result, this study was limited to English speaking respondents.

Overview of the Study

Chapter I. - The Problem provides an introduction to the study through a discussion of the problem, significance of

the study, purpose of the study, research questions, hypotheses tested, assumptions, limitations, and an overview of the dissertation.

Chapter II. - Review of Related Literature discusses concern about training effectiveness, Kaufman's Organizational Elements Model, the "Human Resources Development" concept, and the relative newness of the Organizational Elements Model. A brief summary of these discussions is provided at the end of the chapter.

Chapter III. - Methodologies and Procedures describes the study population, sampling procedure, research design, development of the instrument, data collection, treatment of the data, and analysis procedures. This chapter also is concluded with a brief summary.

Chapter IV. - Discussion of the Results includes discussions of the response rate, reliability of the survey instrument and interviewers, and results. These discussions are briefly summarized at the end of the chapter.

Chapter V. - Summary, Conclusions, and Recommendations includes a summary of the study, conclusions, recommendations for further studies, and recommendations for use of the data presented.

A Bibliography and Appendices are provided at the end of the dissertation.

CHAPTER II

REVIEW OF RELATED LITERATURE

Concern About Training Effectiveness

Concern about training effectiveness has existed for some time. The programmed instruction movement in the 1960's involved the development of methods to empirically determine the effectiveness of programmed instruction (Markle, 1967). By the late 1960's and early 1970's, concepts of both formative and summative evaluation were becoming prevalent in instructional development literature (Lumsdaine, 1965; Scriven, 1967; Stake, 1970; Cunningham, 1971; Baker and Schutz, 1972; Sanders & Cunningham, 1972; Worthen & Sanders, 1973; Popham, 1973). These publications dealt with the question of whether or not instruction is effective in terms of measurable instructional objectives. The models employed concerned the summative judgment of whether or not an acceptable percentage of students master instructional objectives, or the formative determination of what modifications are required to ensure an instructional product or program meets standards for effectiveness.

Formative and summative evaluation methodologies comprised the state of the art for instructional evaluation until very recently. And, formative and summative evaluation methodologies are still important for the determination of the effectiveness and efficiency of instructional products or programs. Increasingly, however, it has been proposed that formative and summative evaluations are not sufficient by themselves to determine the overall worth of training to organizational needs (Gilbert, 1978; Posavac & Carey, 1980; Kearsley, 1982; Fitz-Enz, 1984; Rummler, 1985). Even effective and efficient instructional programs are valueless if they do not contribute to organizational goals, or if they involve content that is counter to organizational goals or client expectations. Therefore, evaluations models which examine the effects of training on organizations are needed in addition to formative and summative evaluation models.

Kaufman's "Organizational Elements" Model

Kaufman (1981, 1983, 1984) has developed a model for organizational needs assessment and training evaluation that utilizes General System Theory concepts (Bertalanffy, 1962 and 1968). Kaufman's model involves examining five elements of organizations that can be used for needs assessment or evaluation (Kaufman & Stone, 1983). These are:

- inputs (resources used)

- processes (the methods to create products)
- products (the result of the process)
- outputs (the result of combining products into a comprehensive whole)
- outcomes (the societal impact of the outputs)

The first four elements (inputs, processes, products, and outputs) are classified by Kaufman as being internal to the organization. For instance in a training program:

- inputs are the resources required to develop and deliver training
- processes are the instructional strategies employed, training environment, instructor behavior, etc.
- products are the learner reaction to instruction, and the learner mastery of instructional objectives
- outputs are the degree to which the learner applies the new behaviors to qualify and perform on the job.

All of these elements (inputs, processes, products, and outputs) are internal to the organization.

The fifth element (outcomes) is, however, external. In the training example, outcomes would be the degree to which learner behavior on the job results in organizational

products, services, or communications that have desired impact on society (increased client satisfaction, profits, or client perception of product value).

The Kaufman model is based upon General System Theory concepts (Bertalanffy, 1968) in that Kaufman views an organization as being a sub-system within a system termed "society." Kaufman maintains that societal factors have to be taken into account because of the relationship of the organization to society, considering the organization as a sub-system of the overall societal system. The outputs of the organization have an impact (outcome) on society in terms of the societal view of the organization. Organizational survival is dependent upon societal impact of the organization. Kaufman, Mayer, and Butz (1984) state that "No organization is a closed, self-sufficient, and self-reliant system. Therefore we must not ignore outcomes."

My interest in the Kaufman model is a result of attending several sessions at the 1985 National Society of Performance and Instruction (NSPI) Conference in which presenters were advocating evaluation of training at the output and outcome organizational elements so that the effect training has on organizational goals can be determined.

It should be noted that several of these presenters were

speaking in terms of Rummler's and Brethower's training evaluation model (Brethower & Rummler, 1979). Rummler and Brethower (Brethower, 1966; Brethower and Rummler, 1979; Rummler, 1985) use flow diagrams to show the points along organizational input/output flows that a training program ought to be evaluated. The points displayed in the Rummler model are in essence the Kaufman organizational elements.

I asked Kaufman and Rummler if their two models are in agreement in terms of organizational elements and both have answered in the affirmative. Rummler uses different terminology, but exact correlations can be found in the Rummler model to the Kaufman Organizational Elements Model. Kaufman and Rummler do differ in their applications of the models: Rummler applies his model at any subsystem level within the organization; Kaufman is primarily interested in the relationship of the organization to society, and only examines organizational sub-system levels in the context of societal impact.

This study would have the same theoretical base if it were done in terms of either model, because both models are applicable to study of organizational impact on society. The Kaufman model was chosen because it can easily be described verbally without the requirement for a flow diagram (although Kaufman often uses a diagram to depict the Organizational Elements Model).

The Kaufman model is fairly recent, being only a few years in print. I believed that it was quite likely that model was not widely represented in terms of training evaluation practice, but no study could be located that indicated the degree of use of the model in evaluation efforts. Given the interest in the Kaufman model in evidence in the 1984 NSPI Conference, it was decided to determine the rationale underlying these practices. More specifically, since it was expected that a large proportion of HRD training programs would not be evaluated in terms of outputs or outcomes.

The "Human Resources Development" Concept

Associated with the requirement for justification of training programs has been the evolution of Personnel departments into integrated Human Resource Development (HRD) departments (Fitz-Enz, 1984). The formation of HRD departments has resulted in the coordination of the traditional functions of employee records and benefits with the newer functions of recruitment and selection, compensation, career development and succession planning, performance appraisal, training, instructional development, and strategic planning. HRD is a concept that has become prevalent throughout most of the large organizations in the United States.

Instructional programs within the context of HRD involve three major types of activities: Training, Education, and Development (Nadler, 1979). Training is carried out to improve job performance (Nadler, 1979, p.40). Education is carried out to improve overall employee competence beyond immediate job requirements (Nadler, 1979, p.60). Development is carried out to prepare the employee for growth opportunities within the organization (Nadler, 1979, p.88). In this study, the term "training" will be used to refer to all three types of activities. For each type of training activity, the major purpose of conducting training is to, "..ensure that skilled manpower is available as needed and that it is optimally productive." (Lusterman, 1977).

The measurement of HRD training function effectiveness requires that HRD training functions be measured in terms of contribution to organizational success. The lack of training program evaluation in terms of organizational success factors can have negative consequences. One danger is that training programs will be aimed at relatively trivial instructional content when compared with organizational need (Mager & Pipe, 1984). Another danger is that even if the training unit is making a strong contribution to organizational success, the lack of

evaluation data may leave the training function vulnerable to staffing reductions without the ability to show the worth of the training outputs.

The Kaufman Organizational Elements Model may be applied to a variety of types of training programs that are offered by HRD training functions. It is possible to classify HRD training programs in a manner different from that described by Nadler, but consistent with the way many HRD departments categorize their training programs; these types include:

- Management development training for managerial and professional staff.
- Job skill training for clerical and technical employees.
- Job orientation programs.
- General health or wellness programs such as learning exercise skills, stress management, or family practice.

Each of these types of programs may involve different training evaluation rationale:

- Management development training programs are often not skill based (although some are). These programs are often evaluated in terms of learner reaction to instruction, and not in terms of learner mastery of performance objectives. It is interesting, however,

that the purpose of management development training is to increase the knowledge and skill (competence) of the managerial force within the organization. It might, therefore, seem highly appropriate to evaluate management development training programs in terms of their effect on organizational outputs. A strong case could even be made for evaluating management development training programs in terms of organizational outcomes, because managers often have strong community ties, and the behavior of managers in meetings with clients or in community service situations could have a profound effect on societal perception of the organization.

- Job skill training is usually performance based, although the training methodologies employed may not be performance based, and often no evaluation of learner mastery of performance objectives is done. It would seem appropriate to evaluate job skill training, however, in terms of organizational products, outputs, and outcomes because the clerical or technical employee plays a major role in organizational productivity. Some trainees may come into direct contact with the clients of the organization, but virtually all employees play a role in helping the organization achieve its production and quality goals.
- Job orientation programs are usually intended for newly hired employees. Orientation programs vary from small

budget informal orientation policies to formal large budget instruction. A major purpose of orientation programs is to ensure that newly hired employees become acclimated to the new working environment and have a productive start in the organization (Latham & Wexley, 1982). Therefore, evaluation of orientation programs in terms of organizational output (employee performance) would seem desirable. Orientation programs are usually not evaluated in terms of employee performance, however, because the programs themselves do not often involve performance objectives. Orientation programs usually center around knowledge objectives (information about the organization and community) and motivational objectives ("See what a good organization you just joined?"). It is not uncommon for orientation programs to be evaluated in only the most cursory way available (the "smiley face" evaluation). Still, if the major purpose of orientations is to make sure that employees get off to productive starts, the degree to which productive starts actually occur should be determined.

- General health or wellness programs are often provided by organizations to improve productivity by encouraging good mental and physical health practices. Such programs can include instruction in stress management, dealing with family problems, exercising, alcohol and drug abuse, smoking avoidance, and community

involvement. These programs are often not evaluated in any other terms than the number of employees that choose to participate in the programs. It would seem appropriate, however, to evaluate wellness programs in terms of the contribution they make to organizational output or outcomes, given that it seems reasonable to justify the expenditure in these programs according to the benefit provided to the organization.

Relative Newness of the Organizational Elements Model

The spread of innovation usually follows a pattern that results in a small percentage of adoption in the early years of the appearance of the innovation (Rogers and Shoemaker, 1971). The Kaufman model represents innovation in evaluation practice in that the Kaufman model, or the similar Rummler model, have been in print for no more than a few years. Thus, according to diffusion of innovation theory, it is likely that only a small percentage of HRD training programs are being evaluated in terms of the Kaufman model at the outputs or outcomes levels. I am highly interested in the extent to which this innovative model has had an effect on current training evaluation practices in view of the relative newness of the model.

Summary

Concern about training effectiveness has existed for a number of years, and was especially evident in the Programmed Instruction movement of the 1960's. In the late 1960's and early 1970's, evaluation concepts developed in the programmed instruction movement were expanded into formative and summative evaluation models. These models represented the state of the art in the 1970's. In recent years, however, several theorists have questioned whether the formative and summative evaluation models by themselves are adequate.

One model aimed at tying together questions of training effectiveness with concepts of organizational impact on society, and organizational health, has been Kaufman's Organizational Elements Model. The Organizational Elements Model involves examining five elements of organizations that can be used for needs assessment or for evaluation. In terms of training programs, the five organizational elements are:

- inputs (the resources required to develop and deliver training)
- processes (the instructional strategies employed, training environment, instructor behavior, etc.)

- products (the learner reaction to instruction, and the learner mastery of instructional objectives)
- outputs (the degree to which the learner applies the new behaviors to qualify and perform on the job).

The Organizational Elements Model is based upon General Systems Theory concepts, in that training programs are viewed as sub-systems of organizations, and organizations are viewed as sub-systems within a system called, "society."

Rummler & Brethower have developed a similar training evaluation model, but use a different set of terminology and support visuals to describe their model. Kaufman and Rummler have agreed with me that their models are interchangeable for the purposes of this study, but the Kaufman model forms the basis of this study because it is more easily described verbally.

The concept of Human Resources Development (HRD) also was an important factor in this study. Traditional Personnel functions in large organizations have evolved into coordinated functions that include training and instructional development. While not all training in organizations is carried out in HRD settings, a large number of HRD departments do have formal training programs. These programs have been classified by Nadler as falling under three categories:

- Training (job knowledge and job skill instruction)
- Education (General improvement of overall competence through instruction)
- Development (instruction aimed at providing employees with growth opportunities within the organization)

I chose, however, to classify training programs according to categories in common use in HRD departments:

- Management development training for managerial and professional employees.
- Job skill training for clerical and technical employees.
- Job orientation programs.
- General health or wellness programs.

Each of the four types of HRD training programs may involve different training evaluation rationale:

- Management development programs should be evaluated in terms of organizational outputs because of the exposure of management personnel in the community, and because of the strong effect management policies have on organizational outcomes.
- Job skill training programs should be evaluated in terms of products, outputs, and outcomes because

clerical and technical employees play a critical role in helping the organization achieve its productivity and quality goals.

- Job orientation programs should be evaluated in terms of outputs and outcomes,in other words; the effect orientation programs have on newly hired employee moral and corporate image.
- Wellness programs should be evaluated in terms of outputs and outcomes given that the reason wellness programs are offered is to bring about a more productive work force through better physical and mental health.

Despite the above reasons of why training programs should be evaluated in terms of outputs and outcomes, the relative newness of the Kaufman model, with the accompanying likelihood that such evaluation practices would not be wide spread (Rogers and Shoemaker, 1971), led me to study the degree to which key concepts of these models have been utilized in HRD training evaluation programs.

CHAPTER III

METHODOLOGIES AND PROCEDURES

Introduction

The methods and procedures used to address the two major research questions of this study are discussed in this chapter. The research questions are:

1. What is the degree of HRD department evaluation practice, in terms of the organizational elements identified by Kaufman, for each of the following types of training program:

   - Orientation
   - Wellness (health practices, exercise, diet, etc.)
   - Job Skill
   - Management Development

   Specifically, what percentage of training programs within each of the above categories are evaluated in terms of products, outcomes, or outputs?

2. For any of the above types of training, if the training

programs are not evaluated in terms of products, outcomes, or outputs, what rationale underlies the lack of such evaluation?

The following sections will outline the methodology to complete the research:

- Study Population
- Sampling Procedure
- Research Design
- Development of the Instrument
- Data Collection
- Treatment of the Data
- Analysis Procedures
- Summary

Study Population

The study population was made up of businesses or organizations within the Association for Supervision and Personnel Administration (ASPA) or the Life Office Management Association (LOMA) having Human Resource Development (HRD) departments. The study population was limited to HRD related training programs because:

- locating managers of training units within line units

of organizations can be a time consuming process, and one can not be sure all training units within a organization have been identified without exhaustive questioning of a number of individuals.

- many training programs exist on an informal basis in line functions; informal training programs are not within the scope of this study.
- of my interest in evaluation of training programs within HRD departments.

The majority of large organizations in the U.S. and Canada have employees holding memberships in ASPA. Similarly, the majority of major insurance organizations (profit or non-profit) in the U.S. and Canada hold memberships in LOMA. According to the 1985 LOMA and ASPA memberships list, the ASPA and LOMA populations total to more than 5,000 individuals.

ASPA and LOMA were chosen because both organizations are well populated with HRD training managers. Other organizations, such as the American Society for Training and Development (ASTD) or the American Management Association also have a large number of members that could be members of this study population. However, the advantage of using ASPA and LOMA membership lists is that their memberships are made up primarily of individuals associated with large organizations (profit or non-profit). A similar membership

profile would also be true of the American Management Association, but much of its membership would be duplicated in the ASPA membership.

Other major reasons for choosing ASPA and LOMA for this study is that their membership lists are sorted by organization as well as by individual, and that ASPA and LOMA memberships represent a solid cross section of HRD departments in the United States and Canada. A high percentage of ASTD's membership is composed of independent training or HRD consultants. Independent consultants are not a part of this study population, and yet they are often listed in the ASTD membership according to consultation firms that have large corporate sounding names, but which may not be large enough to have an HRD department. The ASPA membership list also contains consultation firms, but to a lesser extent than the ASTD membership list.

LOMA is specific to the insurance industry. ASPA and LOMA have recognized that their memberships do not overlap greatly, and these professional groups have cooperated with each other in the past in HRD research projects. A current ASPA/LOMA study is attempting to determine HRD department productivity by surveying more than 100 organizations in the United States and Canada (Fitz-Enz, 1985).

The titles held by people in HRD positions are not

standardized, and are not always descriptive of the duties associated with the positions held. The desired profile of the study participant was that the individual have detailed knowledge (administratively and/or operationally) of the training evaluations conducted in the HRD department within their organization. In order to ensure that the participants in the study matched this profile, each selected institution was contacted by telephone in order to identify and reach persons fitting the profile. Organizations that did not have training functions within the HRD department were considered as being outside the scope of this study.

The names of departments housing HRD functions are also not standardized. Some organizations use the name Personnel department, Corporate Services department, or Administration department. This study considered the function to involve HRD training if it contained any of the four types of training under study, and if the respondent described their job as being in an HRD or Personnel context.

While organizations had to be large enough to have formal HRD training programs, the study population profile did not include criteria for minimum staffing or budgetary levels. The lack of criteria for staffing size or budgeting meant that some organizations included in the study had HRD training functions, but served a small number of employees.

Sampling Procedure

The sample used in this research was randomly drawn from member organizations of LOMA, and ASPA. The procedure for drawing the random sample was as follows:

- Each organization listed in the ASPA and LOMA membership directory was numbered in a way that each organization was numbered only once, and so that no duplicate numbers were assigned.
- More than 300 organizations were selected from the numbered list using a table of random numbers.
- The list was split in half. One list was given to one telephone interviewer, and the other list was given to the other telephone interviewer.
- Each telephone interviewer began making telephone contact with organizations following the randomized order in their list.
- Each telephone interviewer located a member of the study population within each organization. If there was no HRD department in a contacted organization, that organization was taken off the list, and the next organization on the list was contacted.
- Each telephone interviewer continued making contact with organizations until 50 persons (no more than one

person for each organization) had agreed to participate in the study (for a total of 100 participants).

- If the organization was reached, but the participant was not available to answer the interviewer's questions, three follow up calls were attempted.
- If the participant could not be contacted within a total of four telephone calls, that participant's organization was classified as a non-respondent.
- All refusals to participate, or failures to contact participants, were recorded.
- If the number for an organization was not answered on four attempts to call, the number was classified as having a non-respondent.
- Recorded messages that a telephone number was no longer in service, or that a number did not exist, resulted in deleting that name from the eligibility list. Out of service numbers, or non-existing numbers, were considered to be outside the target population.

Research Design

The research design involved surveying participants drawn from a random sample via a structured telephone interview. The random sample design was used because random sampling helps assure representativeness with respect to the population under study. The entire population was not interviewed because telephone interviewing of all members of

the study population would have been prohibitively expensive.

The telephone script (Appendix A.) was used to ensure that persons were informed of who was conducting the survey, why the survey was being conducted, the time the interview might take, the guarantee of anonymity to participants, the type of participant needed, and the fact that there would be no penalty for refusing to participate. Finally, persons were asked if they matched the profile of the participant, or if they could recommend someone. Once contact was established with a person matching the target population profile, the person was asked to participate in the study. If the person agreed to participate the interview began immediately.

Development of the Instrument

The decision to carry out data collection using a telephone survey was made on the following basis (Fowler, 1984):

- Through close supervision, a highly structured telephone survey can result in high reliability.
- Structured telephone surveys provide ease of recording answers in quantifiable form.
- Mailed questionnaires are often answered by people other than the ones intended by the researcher.

Telephone contact helps ensure that the interviewees come from the population under study.

The survey instrument (Appendix B.) was developed using a mixture of closed and open ended questions. The initial plan was to only use closed question design in order to facilitate high reliability and ease of answer tabulation. It became apparent through early evaluation of the survey instrument design however that the questions involving reasons why training programs were not evaluated in certain ways needed to be open ended questions. It is virtually impossible to anticipate all of the reasons that might be given for not conducting evaluations, and presenting respondents with a list of possible reasons would have greatly increased the time involved in conducting a telephone interview. It was decided to simply ask why certain evaluations were not being done, write down the answers verbatim, and classify the answers to these questions later. All other questions consisted of "yes" or "no" answer options, and were therefore instantly classifiable as bivariate data.

It was felt that very few respondents would be familiar with the special terminology used in the Kaufman model. Specific knowledge of the terminology contained in the model was not a part of the desired participant profile. Therefore, the survey instrument was designed using generic training and

evaluation terminology. The survey instrument was reviewed by:

- Lawrence Alexander, Ph.D. (faculty member: Michigan State University),
- Castelle Gentry, Ph.D. (faculty member: Michigan State University),
- Donald Kalmey, Ph.D. (faculty member: University of Louisville), and
- Stephen Yelon, Ph.D. (faculty member: Michigan State University)

to determine the appropriateness of terms used. Their suggestions were implemented in the survey instrument design. In addition, Roger Kaufman reviewed the survey instrument design to ensure that the terminology used in the telephone script and survey instrument was consistent with the constructs contained in his Organizational Elements Model. Kaufman (personal comments, 1986) verified that the terminology was consistent with his model, and offered minor suggestions which were incorporated into the survey instrument design.

A telephone script (Appendix A.) was developed to support the initial telephone contact and screening. Interviewers

were instructed to follow the telephone script closely, and to make sure that all messages contained in the script were communicated to persons completely and accurately.

The initial research plan called for interviewers to be hired from the graduate student population in the School of Education at the University of Louisville; however, Fowler (1985) states that telephone interview reliability is highly dependent upon interviewers, and students used as interviewers require intensive training and supervision in order to ensure interview reliability. It was decided, therefore, to conduct the survey using professional telephone interviewers.

Wilkerson and Associates, a survey company in Louisville, Kentucky was contracted to provide the interviewers and to supervise the interview process. Wilkerson and Associates has a specialized telephone facility that provided low cost telephone connections. In addition, Wilkerson and Associates provides supervisors with capability to listen in on telephone interviews so that interviewer performance can be monitored, and so that improper interview techniques can be identified and corrected. Two interviewers were assigned to this study project. Both interviewers had extensive experience with telephone surveys involving contact with managerial employees in large organizations.

The two interviewers received training that consisted of practice in using the survey instrument and telephone log. In addition, two briefing sessions were held with the interviewers to clarify any questions they had about rules for conducting the interviews.

The survey instrument was pretested by having both interviewers conduct the survey with 10 pretest participants. The pretest participants made prior commitment to the researcher to take two calls for the same telephone survey. Pretest participants were instructed not to try to make their answers consistent, but to try to answer according to how the interviewer presented the questions. Comparison of answers provided revealed that 98% of all answers given were in agreement. A calculation of the Pearson's correlation for the pre-test answers resulted in an inter-rater correlation of .866.

Interviewers were instructed to ask questions exactly as they appeared on the survey instrument. If a respondent asked for clarification of a question, only a repeat of the exact words was allowed. Interviewers were allowed to use minimal probing whenever respondents gave unclear answers to the open ended questions.

Data Collection

Interviewers staggered hours during which telephone calls were made so that no time zone in the United States or Canada received precedence, and so that all persons contacted received their calls in mid morning, or mid afternoon.

Up to three call backs, for a total of four possible telephone calls to each organization, were authorized. If contact could not be made with the desired individual at the organization, the interviewer would ask someone at the number to suggest a time and date to call again. Unanswered rings were recorded as an unsuccessful attempt to call, and attempts were made at least four times to reach numbers that were unanswered.

All 100 respondents were contacted within a three week period in late February and early March, 1986. No attempts were made to call during weekends or holidays.

Interviewers recorded answers on a copy of the survey instrument. A log of telephone calls (Appendix C.) was used to keep track of successful contact, call backs, refusals, and inability to make contact. The first sheet of the survey instrument was also used to record information on the person interviewed. All information that could be used to

identify individuals or organizations was deleted at a later date to protect the anonymity of survey participants.

Answers to closed questions were recorded as simple check marks within the appropriate numbered bracket on the survey instrument. Answers to open ended questions were written verbatim.

Daily meetings were held with interviewers to review filled out telephone logs and survey instrument sheets. Unanticipated problems were discussed in these meetings. Any inconsistencies in data notation were corrected in the first meeting.

Treatment of the Data

Survey instrument sheets were collected daily, and the data on the sheets was transferred into a computer database. The database management program used was Enable, a dBase II compatible program. I personally transferred data from the survey instrument sheets to the computer database. The database file contained fields corresponding to each numbered item on the survey instrument. Check marks on the survey instrument sheets were entered into the database record as a "1". Each field lacking a check mark was entered into the database record as a "0".

Open ended questions did not contain check marks. Instead, the database contained text fields for each open ended question in the survey instrument. Text fields were limited to 78 characters, so that answers to open ended questions were paraphrased for entry into the database.

The contents of the database were printed out in hard copy form, and then compared with the survey instrument sheets to verify the accuracy of data entry. Errors were corrected using the database management system edit facility.

Upon completion of the data entry task, data relating to closed questions was copied from the database into Enable's spreadsheet module, a LOTUS 123 compatible program. Data relating to open ended questions was copied into Enable's word processor using the Enable database management system report module.

Analysis Procedures

An analysis of proportions (Cohen, 1977) was conducted for research hypotheses HO_1 through HO_{16}. These hypotheses were associated with the closed questions in the survey instrument design, and research question 1. A frequency analysis was conducted for research hypotheses HO_{17} through

HO_{19}. These hypotheses were associated with the open ended questions in the survey instrument design, and research question 2.

Analysis of Proportions:

SYSTAT, a statistical analysis program that runs on IBM PC compatible computers, was used to calculate the proportions of respondents reporting that they did specific types of evaluations (learner reaction, learner mastery, outputs, and outcomes) for each of the four types of training programs (orientation, job skill training, management development, and wellness).

Alpha is the probability that the researcher committing a Type I error in hypothesis testing (Cohen, 1977, pp.4-5). The Type I error is an error of rejecting a research hypothesis that is actually true. For testing of HO_1 through HO_{16}, alpha was set a priori at .05. Researchers can also incorrectly accept a research hypothesis (Cohen, 1977, p.5); incorrectly accepting a research hypothesis is known as a Type II or beta error. In general, it is desirable to set both alpha and beta at small levels. Beta for each hypothesis test was calculated following Cohen's procedures.

The Enable spreadsheet module was used to conduct the

hypothesis testing for hypotheses associated with research question 1. I designed a spreadsheet model which computed the effect size, beta, the sampling error, the 67% confidence interval, and the 95% confidence interval for each hypothesis test. Effect size and beta were computed on the basis of power analysis tables published by Cohen (1977, chapter 6.). The spreadsheet model included Cohen's (1977, p.194) power analysis table for two tailed tests of proportions with alpha set a priori at .05. In addition, the spreadsheet included Cohen's (1977, p.183) table for transformation of proportions to arcsine values. The tables in the spreadsheet model were verified for agreement with the published tables.

Sampling error and confidence intervals were computed using formulas published by Fowler (1984, p.36). All formulas and tables were validated by placing data into the spreadsheet model with known outcomes, and verifying that the spreadsheet model produced the desired answers.

An analysis was done at the completion of data gathering to determine the possible effect of non-response upon the validity of the study. The analysis involved estimating that all non-respondents have training programs, but do not evaluate these programs in terms of learner reaction, learner mastery, training outputs, or training outcomes. This very conservative estimate, that no other training

programs are evaluated in any type of way included in the study, was used to compute estimated sample proportions, and to test the research hypotheses associated with research question 1. against these sample proportions.

Frequency Analysis:

I examined printed lists of the reasons given for not carrying out certain types of evaluation, or for not doing any evaluation, for similar responses. The groupings of similar responses were used to make a count of the frequency of each category of response for each open ended question on the survey instrument. The categories of responses were then sorted by order of most frequent category cited by respondents for a given open ended question; the sort was done manually within the word processor. When multiple reasons for not carrying out evaluations were supplied by a respondent, each reason given was counted in the frequency analysis.

Summary

The study population was drawn from ASPA and LOMA membership lists. The two membership lists are complimentary, but any duplications in the lists were eliminated. The population profile called for participants in the study to be members

of HRD departments and to be responsible for, or knowledgeable of, training evaluation practices in their organization.

The sample selection was made through a random selection of organizations, followed by direct telephone contact to identify persons fitting the population profile. Random sampling was used to ensure representativeness of the sample to the target population.

A telephone survey was conducted because:

- telephone surveys using up to date telephone lists typically have higher response rates than mailed surveys,
- a highly structured telephone survey can result in high reliability,
- a structured telephone survey provides ease of recording answers in quantifiable form,
- mailed questionnaires are often answered by people other than the ones intended by the researcher, and
- telephone contact helped to ensure that the interviewees came from the population under study.

A survey instrument, telephone script, and telephone log were developed to facilitate data collection and to ensure reliability of the telephone interviews. Roger Kaufman

reviewed the survey instrument and telephone script for consistency with the constructs in his Organizational Elements Model.

Two professional interviewers were used; the interviewers were under the supervision of Wilkerson and Associates. Wilkerson and Associates provided the telephone facilities and equipment used in the survey.

The survey instrument, telephone script, telephone log, and interviewers were pre-tested. Analysis of the pre-test data revealed a survey instrument consistency of 98% and a inter-rater correlation of .866.

Telephone calls were made across a three week period in February and March, 1986. One hundred persons agreed to participate in the study. All responses were recorded on the survey instrument sheets, and telephone call information was recorded on the telephone log forms. Survey data was stored in a computer database file. Proportions relating to research question 1 were calculated using a statistical analysis program.

Hypotheses relating to research question 1 were tested using a spreadsheet model that was based on Cohen's procedures for power analysis of proportions and Fowler's recommendations for computing sampling errors and confidence intervals.

Hypotheses relating to research question 2 were tested on the basis of frequency analysis, which consisted of grouping the response into like categories, and counting the number of responses in each category.

CHAPTER IV

DISCUSSION OF THE RESULTS

INTRODUCTION

This chapter provides an analysis and discussion of the data collected in the telephone survey. Two major areas of concern are addressed:

1. The proportion of training programs that are being evaluated in terms of learner reaction, learner mastery, training outputs, and outcomes for each of the four types of training programs:
   - orientations,
   - job skill training programs,
   - management development programs, and
   - wellness programs.

2. The most frequently cited reasons for not conducting any types of evaluations, or for not doing output or outcome evaluations.

The first area of concern involves data relating to 16 research hypotheses, and is discussed in terms of an

estimated response rate. An alternative analysis of the data is also discussed. In the alternative analysis, the following assumptions were used:

- All non-respondents conduct training evaluations.
- None of the non-respondents evaluate training programs in terms of types of evaluation of interest in this study.

These assumptions are highly unlikely, and cause the alternative analysis to be based upon very conservative proportion estimates. The alternative analysis was done to demonstrate that all but two hypotheses would have been rejected under the most conservative possible circumstance at alpha=.05.

The second area of concern involves data relating to research hypotheses HO_{17}, HO_{18}, and HO_{19} and is discussed in terms of frequency counts.

RESPONSE RATE

A beginning list of 335 possible participants were selected at random from the LOMA and ASPA membership lists. On the basis of telephone calls, 92 possible participants were screened out because they were identified as not being members of the target population; in other words, 92

possible participants belonged to organizations not having HR departments or lacking training programs to evaluate. The screening out of 92 possible participants left a total potential sample size of 243.

Calls to all 243 potential participants were attempted. 129 of the 243 potential participants were contacted. Of the 129 potential participants contacted, 30 refused to participate in the study, for a response rate of 76%. However, 114 potential participants could not be reached despite four (4) attempted telephone calls to each organization. Persons not reached were also included in response rate calculations because it was not possible to determine if random factors were why these persons could not be reached.

It is reasonable to assume that a portion of those that could not be reached would have been screened out of the study had they been reached. Fowler (1984, pp. 46-47) recommends estimating the percentage that would be screened out according to the percentage of those who were reached that were screened out. Using Fowler's recommendation, it is estimated that 31 of the unreached potential participants would have been screened out of the study. The remaining 83 unreached potential participants were counted as non-respondents. Adding the 83 unreached potential participants to the non-respondent category caused the

estimated response rate to be reduced to 47%. All calculations of sampling error were made on the basis of an estimated 47% response rate.

Initial screening totaled out to 100 respondents. Subsequent analysis of the data showed that one respondent stated at the end of the interview that she didn't have much time, so she answered "yes" to all of the questions. The answers of this respondent were not included in the analysis. This study is therefore based upon 99 participants.

RELIABILITY OF SURVEY INSTRUMENT AND INTERVIEWERS

The telephone script (Appendix A), survey instrument (Appendix B.), and interviewers were pre-tested for consistency by having each interviewer call the same 10 test respondents. Calls from the two interviewers had to be spaced at least 24 hours apart. The test respondents were members of the study target population who agreed in advance to participate in the survey instrument pre-test. These respondents were not included in the actual study.

Test respondents were instructed to make no effort to provide consistent answers to the two interviewers. Instead, test respondents were instructed to provide truthful answers to each question within the context of how the question was asked by the interviewer.

Comparison of the two sets of interview data resulted in a survey instrument consistency rating of 98% (98% of all answers provided to the two interviewers were in agreement). A Pearson's Correlation analysis of the same answers provided an inter-rater correlation of .866.

RESULTS

Proportion of Training Programs Evaluated in Terms of the Model:

All of research hypotheses HO_1 through HO_{16} were rejected (See Table 1., Hypotheses Test Findings). The first 16 research hypotheses were:

HO_1 Less than 5% of HRD orientation training programs are evaluated in terms of learner reaction to instruction.

HO_2 Less than 5% of HRD orientation training programs are evaluated in terms of learner mastery of objectives.

HO_3 Less than 1% of HRD orientation training programs are evaluated in terms of outputs.

HO_4 Less than 1% of HRD orientation training programs are evaluated in terms of outcomes.

HO_5 Less than 20% of HRD job skill training programs are evaluated in terms of learner reaction to instruction.

HO_6 Less than 10% of HRD job skill training programs are

evaluated in terms of learner mastery of objectives.

HO_7 Less than 1% of HRD job skill training programs are evaluated in terms of outputs.

HO_8 Less than 1% of HRD job skill training programs are evaluated in terms of outcomes.

HO_9 Less than 20% of HRD management development training programs are evaluated in terms of learner reaction to instruction.

HO_{10} Less than 1% of HRD management development training programs are evaluated in terms of learner mastery of objectives.

HO_{11} Less than 1% of HRD management development training programs are evaluated in terms of outputs.

HO_{12} Less than 1% of HRD management development training programs are evaluated in terms of outcomes.

HO_{13} Less than 5% of HRD wellness training programs are evaluated in terms of learner reaction to instruction.

HO_{14} Less than 1% of HRD wellness training programs are evaluated in terms of learner mastery of objectives.

HO_{15} Less than 1% of HRD wellness training programs are evaluated in terms of outputs.

HO_{16} Less than 1% of HRD wellness training programs are evaluated in terms of outcomes.

In each case alpha was set a priori at .05 for a two tailed test of proportions. The statistical tests were based upon power analysis procedures specified by Cohen (1977, pp.

179-213). Table 1. also indicates the sampling error, 67% confidence interval, and 95% confidence interval for each percentage finding.

Table 1. Hypotheses Test Findings

Hypotheses	n	Hypothesized Proportion	Sample Proportion*	Sampling Error	67% Confidence Interval	95% Confidence Interval
HO_1	63	0.05	0.59	0.06	0.53-0.65	0.47-0.71
HO_2	63	0.05	0.49	0.06	0.43-0.55	0.37-0.61
HO_3	63	0.01	0.49	0.06	0.43-0.55	0.37-0.61
HO_4	63	0.01	0.24	0.05	0.19-0.29	0.14-0.34
HO_5	44	0.20	0.93	0.04	0.89-0.97	0.85-1.00
HO_6	44	0.10	0.84	0.06	0.78-0.90	0.72-0.96
HO_7	44	0.01	0.86	0.05	0.81-0.91	0.76-0.96
HO_8	44	0.01	0.36	0.07	0.29-0.43	0.22-0.50
HO_9	56	0.20	0.93	0.03	0.90-0.96	0.87-0.99
HO_{10}	56	0.01	0.70	0.06	0.64-0.76	0.58-0.82
HO_{11}	56	0.01	0.61	0.07	0.54-0.68	0.47-0.75
HO_{12}	56	0.01	0.41	0.07	0.34-0.48	0.27-0.55
HO_{13}	22	0.05	0.46	0.11	0.35-0.57	0.24-0.68
HO_{14}	22	0.01	0.55	0.11	0.44-0.66	0.33-0.77
HO_{15}	22	0.01	0.55	0.11	0.44-0.66	0.33-0.77
HO_{16}	22	0.01	0.27	0.09	0.18-0.36	0.09-0.45

* All research hypotheses rejected at alpha of .05.

Table 1.

Of the 99 qualified participants that agreed to take part in the survey, 33 reported they did not do any type of training evaluation. For these 33 respondents, no data was gathered as to the type of training programs their departments offered. Table 1., therefore, shows the evaluation percentages associated with departments that run some sort of training evaluation.

For each hypothesis test, beta was at the .005 level. Beta

of .005 is a high level of statistical power, and indicates that there was less than a .5% chance of failing to reject a false research hypothesis, or that there is a 99.5% chance that the researcher correctly rejected false research hypotheses.

Additional tests were conducted of HO_1 through HO_{16} with the assumption that the 33 respondents not conducting training evaluations also had the same percentage of orientation, job skill training, management development, and wellness programs as those respondents who did carry out evaluations. The assumption of similar proportions of types of training programs caused the estimated number of training programs to go up to 84, 59, 74, and 29 respectively. With these estimated sample cell sizes, estimated sample proportions were reduced, but all research hypotheses would still have been rejected with alpha set a priori at .05. Table 2., Estimated Sample Proportion Tests, indicates what the sample proportions would look like if all calls could have been completed, and if all of these calls resulted in answers of:

- We have all four types of programs,
- We do some training evaluation,
- We do not evaluate any training program for learner reaction,
- We do not evaluate any training program for learner mastery,

- We do not evaluate any training program for outputs, and
- We do not evaluate any training program for outcomes.

It is highly unlikely that none of the training evaluations are of the four types listed in Table 2., but making calculations using these assumptions provides a very conservative estimate of what the true sample proportions are. Table 2. shows that using these very conservative estimates, all but two of the research hypotheses, HO_5 and HO_{13}, would have been rejected at the .05 level.

Table 2., Estimated Sample Proportion Tests

Hypotheses	n	Hypothesized Proportion	Estimated Sample Proportion	Sampling Decision*	Error
HO_1	202	0.05	0.18	Reject	0.03
HO_2	202	0.05	0.15	Reject	0.03
HO_3	202	0.01	0.15	Reject	0.03
HO_4	202	0.01	0.07	Reject	0.02
HO_5	141	0.20	0.29	Accept	0.04
HO_6	141	0.10	0.26	Reject	0.04
HO_7	141	0.01	0.27	Reject	0.04
HO_8	141	0.01	0.11	Reject	0.03
HO_9	180	0.20	0.29	Reject	0.03
HO_{10}	180	0.01	0.22	Reject	0.03
HO_{11}	180	0.01	0.19	Reject	0.03
HO_{12}	180	0.01	0.13	Reject	0.03
HO_{13}	71	0.05	0.14	Accept	0.04
HO_{14}	71	0.01	0.17	Reject	0.05
HO_{15}	71	0.01	0.17	Reject	0.05
HO_{16}	71	0.01	0.08	Reject	0.03

* alpha = .05

Table 2.

Reasons Why Training Programs Are Not Evaluated (HO_{17}):

HO_{17} The most frequently given reason for not evaluating any type of training program in terms of learner mastery is "Costs too much."

HO_{17} is rejected. The given reason that training evaluation is not needed was provided as frequently as the given reason that resources (time, money, staff) were lacking to carry out training evaluations.

33% of the 99 respondents reported that no training evaluations were being conducted. Respondent's reasons for not conducting evaluations are listed in Appendix D. The reasons given can be classified into 11 different categories (the number of responses providing that type of reason is shown in parentheses):

- We don't need to evaluate, we know where we stand (6)
- Not enough resources (time, staff, money) are available (6)
- Plan to start an evaluation program (3)
- Our program is too small to worry about formal evaluation (3)
- We don't run formal training programs, so don't need formal evaluation (3)
- Our training program is too new (3)

- We only need informal evaluations (3)
- Don't know how to evaluate training, or it's hard to measure (3)
- Evaluation done by parent company or consultant (2)
- We contract the training out and let the contractor worry about evaluation (1)

Reasons Why Training Programs Are Not Evaluated for Outputs (HO_{18}):

HO_{18} The most frequently given reason for not evaluating any type of training program in terms of output is "Would like to, but not enough budget."

HO_{18} is rejected. The most frequently given reasons for not evaluating training in terms of outputs were that informal evaluations are sufficient, or that it is not known why output evaluation is not done (8 cases each). Difficulty of measurement was given as the reason in 7 cases. Budget was not mentioned at all, and lack of resources was mentioned by only 4 respondents.

51% of orientation programs, 14% of job skill training programs, 39% of management development programs, and 45% of wellness programs are not evaluated in terms of outputs (on the job behavior). The respondent's reasons for not collecting data on outputs are listed in Appendix D. The

reasons given can be classified into 12 different categories (the number of responses providing that type of reason is shown in parentheses):

- We only need informal evaluations (8)
- Don't know why we don't evaluate outputs, or no reason is provided (8)
- Don't know how to evaluate training, or it's hard to measure (7)
- We don't need to evaluate, we know where we stand (5)
- We don't have the time (4)
- Not enough resources are available (4)
- Plan to start an evaluation program (2)
- We contract the training out and let the contractor worry about evaluation (1)
- Our training program is too new (1)
- Lack of management interest in evaluation of outputs (1)
- We don't have the authority to evaluate on the basis of job behavior (1)
- Training is compulsory, so we don't evaluate in relation to job behavior (1)

Reasons Why Training Programs Are Not Evaluated for Outcomes (HO_{19}):

HO_{19} The most frequently given reason for not evaluating

any type of training program in terms of outcomes is "Can't be measured."

HO_{19} is rejected. The most frequently given reason for not evaluating training in terms of outcomes were that outcome evaluation is not needed. Difficulty of measurement was the second most frequently given reason for not evaluating training in terms of organizational outcomes.

Respondents reported that 76% of orientation programs, 64% of job skill training programs, 59% of management development programs, and 73% of wellness programs are not evaluated in terms of outcomes (societal impact of training programs). The respondent's reasons for not collecting data on outcomes is listed in Appendix D. The reasons given can be classified into 12 different categories shown in the order of the most often given reasons (the number of responses providing that type of reason is shown in parentheses):

- We don't need to evaluate, we know where we stand (28)
- Don't know how to evaluate training, or it's hard to measure (27)
- Our training program is too new (12)
- Not enough resources are available (7)
- Someone else is responsible for evaluation of outcomes (6)

- Don't know why we don't evaluate this (4)
- We haven't ever thought of doing evaluation of outcomes (3)
- We can't act on the results of outcome evaluations (3)
- We don't have the time (3)
- Lack of management interest in evaluation of outcomes (2)
- We're too small an organization to do evaluation of outcomes (2)
- We only need informal evaluations (1)

Summary

Data was collected in the telephone survey and analyzed to address two major areas of concern:

1. The percentage of training programs that are being evaluated in terms of learner reaction, learner mastery, training outputs, and outcomes for each of the four types of training programs: orientations, job skill training programs, management development programs, and wellness programs.

2. The most frequently cited reasons for not conducting any types of evaluations, or for not doing output or outcome evaluations.

335 persons were selected at random from the LOMA and ASPA membership lists as potential survey participants. Of the 335, 92 were contacted and screened out of the study. 100 agreed to participate, with one of these screened out on subsequent analysis of the data collection sheets. 30 contacted persons refused to participate. 114 persons could not be contacted. It was estimated that 31 of the 114 not contacted would have been screened out of the study. The estimated response rate was 47%.

For the hypotheses related to concern area 1., a two tailed test of proportions, with alpha set a priori at .05, resulted in rejection of all research hypotheses, HO_1 through HO_{16}. A recalculation, using a conservative estimate that all uncontacted participants would have reported no evaluations for each type of training program, would have resulted in rejection of all but two of the hypotheses related to concern 1: HO_5 and HO_{13} with alpha set a priori at .05.

An analysis of frequencies resulted in rejection of the hypotheses related to concern 2: HO_{17}, HO_{18}, and HO_{19}. Namely, the most frequently cited reasons for not conducting any types of evaluations, or for not doing output or outcome evaluations.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

SUMMARY

The findings of this study indicate that evaluation practices at the learner reaction, learner mastery, output, and outcome levels are not uncommon. Training evaluation practices within HRD departments appear to be much more consistent with the Kaufman Organizational Elements Model than I hypothesized. It seems that the idea of evaluating training programs in terms of effect on behavior on the job, and in terms of societal impact have become accepted in a significant (practically and statistically) number of Human Resource departments.

CONCLUSIONS

Conclusion 1: Training evaluation practices, in terms of the Kaufman Organizational Elements Model, are much more common than hypothesized.

Despite the low response rate of 47%, the findings make it clear that a significant proportion, in terms of both statistical and practical significance, of HRD training

programs are evaluated in ways that are consistent with the Kaufman Organizational Elements Model. That so many programs are evaluated in terms of the Kaufman Organizational Elements Model is in opposition to my previous belief, and in opposition to the beliefs of both Kaufman and Rummler. When presented with these findings, Geary Rummler commented that he could not argue with the survey methodology, but that he still had difficulty believing the data. He said one possible source of invalidity might simply be that respondents did not tell the truth. Roger Kaufman (personal comment, 1986) also expressed surprise at the high proportion of data collection at the output and outcome levels found.

A possible explanation for my , and the model theorists, large underestimation of output and outcome evaluations is that proponents of the models may place too much weight on instances of non-evaluation in their experience. It is entirely possible that evaluation practices have spread more rapidly than the theorists expected, and that current concern in the United States and Canada over issues of productivity, quality, and competition have led many evaluation specialists to link their evaluation programs into measures of organizational outputs and outcomes.

Conclusion 2: Reasons for not carrying out training evaluations are most often centered around a lack of interest in training, or because of cost or resource constraints.

Both reasons for not evaluating training were cited an equal number of times by respondents. However, the small sample size of 33 calls into question the representativeness of the sample to the population of HRD training programs that are not evaluated. No single reason for not evaluating training was provided by a majority of the respondents. Caution must be exercised in generalizing Conclusion 2 to the target population.

Conclusion 3: Reasons for not collecting data on training outputs (on the job behavior related to training) are most often centered around unperceived need for formal evaluation of training outputs, or that the idea of conducting evaluation of outputs had not even occurred to the person responsible for training evaluation.

By grouping all four types of training programs together, common reasons for not collecting data on training outputs did appear. However, some of the multiple reasons were due to the same respondent giving the same reason for each of

several types of training programs. Again, the representativeness of the sample should be questioned because of the very small sample size, and because of the unequal weighting of responses (not all respondents reported reasons for more than one type of training program). It is plausible, however, that training output data is not collected because evaluators have not reached an awareness stage, or an acceptance stage, of the value of training output data. This explanation, that evaluators are not yet at awareness or acceptance stages for evaluation of training output, is consistent with theories of the diffusion of innovations (Rogers and Shoemaker, 1971) considering the newness of the Organizational Elements Model.

Conclusion 4: Reasons for not collecting data on training outcomes (societal impact of training) are most often centered around a perception that evaluation of outcomes is not needed, or that training outcomes are difficult to measure.

All of the cautions associated with Conclusion 3. apply to Conclusion 4. However, more than 25% of the reasons given involved expression of lack of need to evaluate. Similarly, more than 25% of the reasons given involved lack of knowledge about how to measure training outcomes. It would be reasonable to assume that both types of reasons for not collecting data on outcomes are prevalent throughout the

target population, although the actual percentages may vary greatly from those shown in this study.

RECOMMENDATIONS

Recommendations for Further Studies:

Given that the proportion findings were far above the intuitive beliefs of Kaufman, Rummler, and myself, and considering the possible invalidity introduced through the low 47% response rate, the following recommendation for further studies is offered:

(1) That this study be replicated with a budget large enough to ensure a response rate above 70%. Replication of this study is necessary not only to ensure repeatability of findings, but also to correct the low response rate achieved in the initial study. Fowler (1984) states that any survey with a response rate of less than 70% potentially lacks validity.

I limited the study to HRD training programs. Such programs only represent a portion of organizational training programs in the United States and Canada, and evaluation practices in other types of training environments may be radically different. The following recommendations are offered:

(2) Investigate training evaluation practices in other training environments. It is possible that training evaluation practices vary by type of industry, or by type of department in which training programs are administered. Future studies could examine the adoption of evaluation practices on an industry specific basis, and could determine if certain types of training environments are specially receptive or resistant to newly developed evaluation models.

(3) Study the quality of evaluation efforts. This study did not determine the quality of data collection efforts, or the quality of evaluation decisions made by respondents. It could be that even though a large proportion of evaluators are collecting product, output, and outcome data, that they are not using proper data collection techniques, or that they are not properly analyzing the data. Rather than conduct a large scale survey, an in depth study of selected training evaluation programs could uncover typical problems with data collection in the areas of products, outputs, and outcomes, and could reveal techniques for measuring training outputs and outcomes in a valid manner.

(4) Study the diffusion and adoption of the Organizational Elements Model in a large organization. Ethnographic

study techniques could be used to provide rich information on the problems encountered in introducing the Organizational Elements Model into practice. Accounts could be written of strategies employed: to communicate the value of the model to non-training management; to gain authorization for funding the evaluation program; to meet the resistance of management and employees to allowing the training function access to job performance and organizational performance data; for analyzing the data; and for reporting conclusions in a meaningful manner.

(5) Study the effect of organizational size on training evaluation practices. Smaller training programs may find it easier to collect data due to ready access to information on performance. For instance, if the training audience is small, and not geographically spread, it might be possible to directly observe job behavior to determine the effect of training on job performance. Smaller organizations also might find it easier to gather information from clients, or to check the quality of outputs. On the other hand, smaller organizations may lack the resources to carry out proper evaluations. It would be valuable to know if there is a size advantage for large organizations, and

if so what the critical size and resource factors are that contribute to, or inhibit, the implementation of the Organizational Elements Model.

(6) Study the possible difference in training evaluation practice in unionized and non-unionized manufacturing and service industries. Such a study would help managers determine the possible effect of unionization on facilitation of, or resistance to, transfer of learned behaviors onto the job, or the effect of training programs on organizational outputs and outcomes.

(7) Pursue collection of carefully thought out explanations for not conducting training product, output, or outcome evaluations. Use a structured interview format to determine the variety of factors which may contribute to the decision not to evaluate, and have respondents explain their rationale by providing examples of resistance they have encountered, reasons why they themselves don't believe in the value of specific evaluation practices, or the most important constraints on their evaluation efforts.

Recommendations for Use of the Data Presented:

Recommendations are offered in view of the following limitations of this study:

- The methods used by participants to collect evaluation data were not identified,
- Sample sizes ranging from 22 to 63 led to sample errors from .03 to .11,
- A 47% response rate, and
- The study was limited to English speaking respondents.

Specific recommendations:

(1) Limit generalizations to HRD training programs. Training programs housed in line function departments may have radically different evaluation practices. For instance, "buyer's studies" are commonplace in the insurance industry to determine the receptiveness of customers to product offerings, to sales approaches, etc. Sales training functions in the insurance industry may have ready access to such organizational outcome data, and may already practice evaluating training effectiveness on the basis of information gained from "buyer's studies." Conversely, regulated and monopoly industries may feel that they have a captive audience in their clientele, and may feel little worry about organizational outcomes due to lack of competition.

(2) Training managers and evaluation professionals should critically examine the reasons given for not evaluating, to determine the validity of the arguments. Given the expense of training, is non-evaluation of training effectiveness justified in any instance? The answer may be that in some instances formal evaluation does not lead to changes in training programs, or that the need for training is obvious and the determination of effectiveness is easy. But, should evaluation be avoided simply because it can be expensive or difficult?

(3) Examine higher education curricula to determine if the constructs of the Organizational Elements Model should be taught to program evaluation students and to instructional development students. Given the importance of training outputs and outcomes to organizational performance, students of evaluation should be familiar with a variety of training evaluation models. Student familiarity should include understanding the purpose of the models, the potential utility of models, the sources of resistance to application of the models, and strategies for advocating the use of the models. Use the collection of reasons given for not collecting product, output, or outcome data to form part of the knowledge base for the curricula.

BIBLIOGRAPHY

Baker, R.L. and Schutz, R.E. (Eds.). (1972). Instructional Product Research. Cincinnati, Ohio: Van Nostran Reinhold Co.

Bertalanffy, L. (1968). General System Theory. New York: George Braziller Publishers.

Bertalanffy, L. (1962). General System Theory: A Critical Review. General Systems, VII., 1-20.

Brethower, D.M. (1966). A Systems Approach to Educational Problems. Occasional Paper, Center for Programmed Learning for Business, University of Michigan.

Brethower, K.S. & Rummler, G.A. (1979, May). Evaluating Training. Training and Development Journal.

Cohen, A. & Hyman, J.S., (1979, December). How Come So Many Hypotheses in Educational Research are Supported? Educational Research, 12-16.

Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences. (Rev. Ed.) New York: Academic Press.

Corrigan, G.J. (1980, January). Corporate Training: A Career for Teachers? Phi Delta Kappan, 61(5), 328-331.

Cunningham, D.J. (1971, February) Formative Evaluation of Replicable Forms of Instruction. Paper presented at the annual meeting of the American Educational Research Association, New York. Fitz-Enz, J. (1984). How to Measure Human Resources Management. N.Y.: McGraw-Hill, Inc.

Fitz-Enz, J. (1985). Human Resource Effectiveness Survey: 1st Quarter 1985. Saratoga, Calif.: Saratoga Institute. Fowler, F.J. Jr. (1984). Survey Research Methods. Beverly Hills, California: Sage Publications, Inc.

Gael, S. (1983). Job Analysis. San Francisco, Calif.: Josey-Bass. Gilbert, T.F. (1978). Human Competence: Engineering Worthy Performance. New York: McGraw-Hill.

Hays, W.L. (1981). Statistics. New York: Holt-Rinehart & Winston.

Horrigan, J.T. (1979, July). The Effects of Training on Turnover: A Cost Justification Model. Training and Development Journal, 3-8.

Kaufman, R. and Stone, B. (1983). Planning for Organizational Success. New York: John Wiley & Sons.

Kaufman, R. and Mayer, H. (1981, February). Educating for Beyond Performance and Instruction: A Pragmatic Appraisal. NSPI Journal, 8-12.

Kaufman, R., Mayer, H., and Butz, R. (1984, May). Defining and Classifying the Organizational Elements. Performance and Instruction Journal, 22-25.

Kearsley, G. (1982). Costs, Benefits, & Productivity in Training Systems. Reading, Mass.: Addison-Wesley.

Latham, G.P. & Wexley, K.N. (1982). Increasing Productivity Through Performance Appraisal. Reading, Mass.: Addison-Wesley.

Lumsdaine, A.A. (1965). Assessing the Effectiveness of Instructional Programs. in Glaser (Ed.) Teaching Machines and Programmed Learning. Washington, D.C.: NEA.

Lusterman, S. (1977). Education in Industry. New York: The Conference Board.

Mager, R.F. (1984). Measuring Instructional Results. Belmont, California: Pitman Learning, Inc.

Mager, R.F. & Pipe, P. (1984). Analyzing Performance Problems. (2nd Edition) Belmont, California: Pitman Learning, Inc.

Markle, S.M. (1967). Empirical Testing of Programs. Sixty-Sixth Yearbook of the National Society for the Study of Education, Part II, 104-138.

Nadler, Leonard. (1979). Developing Human Resources, 2nd ed. Austin: Learning Concepts.

Nie, N.H., Hull, G.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. (1975). Statistical Package for the Social Sciences, 2nd Ed.

Popham, W.J. (1973). Evaluating Instruction. Englewood Cliffs, N.J.: Prentice-Hall.

Posavac, E.J. and Carey, R.G. (1980). Program Evaluation: Methods and Case Studies. Englewood Cliffs, N.J.: Prentice-Hall.

Rogers, E.M. and Shoemaker, F.F. (1971). Communication of Innovations: A Cross-Cultural Approach. N.Y.: Free Press.

Rummler, G.A., (1985). Organizations as Systems: Introduction and Basic Premise. Summit, New Jersey: The Rummler Group.

Sanders, J.R. and Cunningham, D.J. (1972). A Structure for Formative Evaluation in Product Development. Educational Research and Evaluation Laboratory, Indiana University.

Scriven, M. (1967). The Methodology of Evaluation. AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally and Co., 39-83.

Shipp, T. (1980, November). To Survive the Budget Inquisition, Prove Your Training Makes Dollars and Sense. Training/HRD, 23-29.

Sibson, R.E. (1985). The Management of Personnel. N.Y.: Vantage Press.

Stake, R.E. (1970). Objectives, Priorities, and Other Judgment Data. Review of Educational Research, 40, 181-212.

Thompson, M.S. (1980). Benefit-Cost Analysis for Program Evaluation. Beverly Hills, Calif.: Sage.

Tyler, R.W. (1931). What is Statistical Significance? Educational Research Bulletin, 10, 115-118.

Wilkensen, G.L. (1972). Cost Evaluation of Instructional Strategies. AVCR, 21.

Worthen, B.R. and Sanders, J.R. (1973). Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Co.

APPENDIX A

Telephone Script

Telephone Script & Survey Instructions

THE INITIAL CONTACT

"My name is (your name). I am trying to reach _____ ."

If not in, "Would you please take a message?" , or

"Could you suggest a time for me to call back?"

If in, "Hello, _____, my name is (your name), and I am calling you concerning a survey being conducted for a doctoral dissertation by Duane Stevens of Michigan State University. The survey is on training evaluation practices."

NOTE: If at any time the person wishes to speak to Duane Stevens, that is allowed.

Take a message, and Duane will call the person.

PURPOSE

"The purpose of the survey is to determine what kinds of training evaluations are done by Human Resources Departments, and to find out the reasons why certain types of evaluations aren't done. This is being done to find out what the state of the art is in terms of Human Resources training evaluation."

TYPE OF SURVEY

"The survey is being conducted entirely by telephone to save participant time. It takes less than 10 minutes to answer the questions."

TYPE OF PARTICIPANT NEEDED

"The person answering the questions needs to be in your company's Human Resources or Personnel Department, and needs to know the Human Resources or Personnel Department's training evaluation practices. Does this describe you, or should I be talking to someone else?"

If someone else: "Would you please give me that person's name and telephone number?"

GETTING COMMITMENT

"I will ask you to be a participant in the survey, but first I will explain some things you should know about participating in this survey." "It is important for you to know that your answers will be held in strict confidence. No one except the researcher will know who you are, or what company you represent. You, or your company, will not be identified in any publication, presentation, or conversation. Your answers will be combined with other people's answers, and will be reported as summary information only. In other words, your anonymity is guaranteed."

"There is no penalty for not participating. The researcher will never mention you to anyone even if you decide not to participate."

"The researcher will keep a file of names and addresses so that survey results can be mailed to participants. That is why I will ask you, at the end of the survey, for your full name and mailing address. If you don't want to receive the survey results, then I won't ask you for your mailing address. At any rate, the address file will be destroyed as soon as the survey results are mailed to you and the other participants."

Do you have any questions?"

"If there are no (more) questions, will you be a participant in this study?"

If, "No", thank the person for talking with you. Write "Answered, 'No'" in the Results section of the Telephone Log.

If, "Yes" - say, "As I said, the interview will take only about 10 minutes, can we conduct the interview right now?"

CONDUCTING THE SURVEY

Start by explaining that the survey will involve a repetitive pattern of questions, that you are not allowed to change the pattern, but that it won't take much time to go through all of the questions. Explain that the repetitive pattern to the questions is important to the research design.

Ask the survey questions on a word for word basis. If a participant doesn't understand a question, take notes on your explanation for the researcher to view later. At the end of the survey thank the participant and promise to send the survey results if requested by the participant.

APPENDIX B

Survey Instrument

SURVEY INSTRUMENT

Company Number_____(1)

Company Name _____(2)

Profit or Non-Profit Organization?_____(2a)

Type of Business or Service?_____(2b)

Name of Participant _____(3)

Telephone Number _____(4)

Date _____(5)

Time _____(6)

Does your department evaluate the effectiveness of your training programs?

Yes - Go to Page 1. (7)

No - Why don't you evaluate your training programs? (8-18)

_____

_____

_____

_____

_____

_____

Go to Page 13 if the answer was "No."

This page intentionally left blank.

Page 1 Starts on the Next Page

I. Does your department offer management development training programs (In other words, performance appraisal, time management, interpersonal skills, etc.)?

Yes ___(19) Go to I.A. No ___(20) Go to Page 4.

I.A. Do you evaluate your management development training programs?

Yes ___(21) Go to I.A.1. No ___(22) Go to Page 4.

I.A.1. Area of evaluation:

Do you collect data on training costs or inputs, in other words facilities, trainers, materials, trainee salary, lost time from the job?

Yes _____ (23) No _____ (24)

Do you collect data on processes such as training method, instructor abilities, training environment, or trainer behavior?

Yes _____ (25) No _____ (26)

Do you collect data on learner reaction to instruction; for example, end of session questionnaires or surveys?

Yes _____ (27) No _____ (28)

If "No": Why don't you evaluate the learner's reaction to instruction for your management development training programs? (29-39)

_____

_____

_____

_____

_____

_____

_____

Do you collect data on learner mastery of training objectives?

Yes _____ (40) No _____ (41)

If "No": Why don't you evaluate the learner's mastery of training objectives for your management development training programs? (42-52)

_____

_____

_____

_____

_____

_____

_____

_____

_____

Do you collect data on learner behavior on the job?

Yes _____ (53) No _____ (54)

If "No": Why don't you evaluate the learner's behavior on the job relating to your management development training programs? (55-65)

_____

_____

_____

_____

_____

_____

Do you collect data on such things as how this type of training eventually effects the outcomes outside of your company in terms of client satisfaction, attitudes towards your company, attitudes towards your company's products, profitability, repeat sales, customer increase, etc.?

Yes _____ (66) No _____ (67)

If "No": Why don't you evaluate the effects your management development training programs have on these kinds of corporate outcomes? (68-78)

II. Does your department offer job skill training programs for clerical, technical, or professional employees?

Yes ___(79) Go to II.A. No ___(80) Go to Page 7.

II.A. Do you evaluate your job skill training programs?

Yes ___(81) Go to II.A.1 No ___(82) Go to Page 7.

II.A.1. Area of evaluation:

Do you collect data on training costs or inputs, in other words facilities, trainers, materials, trainee salary, lost time from the job?

Yes _____ (83) No _____ (84)

Do you collect data on processes such as training method, instructor abilities, training environment, or trainer behavior?

Yes _____ (85) No _____ (86)

Do you collect data on learner reaction to instruction; for example, end of session questionnaires or surveys?

Yes _____ (87) No _____ (88)

If "No": Why don't you evaluate the learner's reaction to instruction for your job skill training programs? (89-99)

_____

_____

_____

_____

_____

_____

_____

Do you collect data on learner mastery of training objectives?

Yes _____ (100) No _____ (101)

If "No": Why don't you evaluate the learner's mastery of training objectives for your job skill training programs? (102-112)

Do you collect data on learner behavior on the job?

Yes _____ (113) No _____ (114)

If "No": Why don't you evaluate the learner's behavior on the job relating to your job skill training programs? (115-125)

Do you collect data on such things as how this type of training eventually effects the outcomes outside of your company in terms of client satisfaction, attitudes towards your company, attitudes towards your company's products, profitability, repeat sales, customer increase, etc.?

Yes _____ (126) No _____ (127)

If "No": Why don't you evaluate the effects your job skill training programs have on these kinds of corporate outcomes? (128-138)

III. Does your department offer training programs in the areas of health, family services, diet....that sort of thing?

Yes ___(139) Go to III.A. No ___(140) Go to Page 10.

III.A. Do you evaluate this sort of training program?

Yes ___(141) Go to III.A.1. No ___(142) Go to Page 10.

III.A.1. Area of evaluation:

Do you collect data on training costs or inputs, in other words facilities, trainers, materials, trainee salary, lost time from the job?

Yes _____ (143) No _____ (144)

Do you collect data on processes such as training method, instructor abilities, training environment, or trainer behavior?

Yes _____ (145) No _____ (146)

Do you collect data on learner reaction to instruction; for example, end of session questionnaires or surveys?

Yes _____ (147) No _____ (148)

If "No": Why don't you evaluate the learner's reaction to instruction for this type of training program? (149-159)

_____

_____

_____

_____

_____

_____

_____

_____

Do you collect data on learner mastery of training objectives?

Yes _____ (160) No _____ (161)

If "No": Why don't you evaluate the learner's mastery of training objectives for this type of training program? (162-172)

_____

_____

_____

_____

_____

_____

_____

_____

_____

Do you collect data on learner behavior on the job?

Yes _____ (173) No _____ (174)

If "No": Why don't you evaluate the learner's behavior on the job relating to this type of training program? (175-185)

_____

_____

_____

_____

_____

_____

_____

Do you collect data on such things as how this type of training eventually effects the outcomes outside of your company in terms of client satisfaction, attitudes towards your company, attitudes towards your company's products, profitability, repeat sales, customer increase, etc.?

Yes _____ (186) No _____ (187)

If "No": Why don't you evaluate the effects this type of training program has on these kinds of corporate outcomes? (188-198)

IV. Does your department offer orientation programs for newly hired employees?

Yes ___(199) Go to IV.A. No ___(200) Go to Page 13.

IV.A. Do you evaluate your orientation programs?

Yes ___(201) Go to IV.A.1. No ___(202) Go to Page 13.

IV.A.1. Area of evaluation:

Do you collect data on training costs or inputs, in other words facilities, trainers, materials, trainee salary, lost time from the job?

Yes _____ (203) No _____ (204)

Do you collect data on processes such as training method, instructor abilities, training environment, or trainer behavior?

Yes _____ (205) No _____ (206)

Do you collect data on learner reaction to instruction; for example, end of session questionnaires or surveys?

Yes _____ (207) No _____ (208)

If "No": Why don't you evaluate the learner's reaction to instruction for orientation programs? (209-219)

_____

_____

_____

_____

_____

_____

Do you collect data on learner mastery of training objectives?

Yes _____ (220) No _____ (221)

If "No": Why don't you evaluate the learner's mastery of training objectives for orientation programs? (222-232)

Do you collect data on learner behavior on the job?

Yes _____ (233) No _____ (234)

If "No": Why don't you evaluate the learner's behavior on the job relating to orientation programs? (235-245)

Do you collect data on such things as how orientation programs eventually effect the outcomes outside of your company in terms of client satisfaction, attitudes towards your company, attitudes towards your company's products, profitability, repeat sales, customer increase, etc.?

Yes _____ (246) No _____ (247)

If "No": Why don't you evaluate the effects orientation programs have on these kinds of corporate outcomes? (248-258)

Thank you very much for your help with this survey. The results of the survey will be available several months from now. We will provide you with a summary report of the results if you would like.

Would you like to receive a summary report?

Yes - Please give me your:

Title:_____(259)

Address:_____ (260)

_____ (261)

City:_____ (262)

State:_____(263) Zip Code:____ (264)

Thank you again for your help. You, and your company, will be kept anonymous in any report associated with this interview. Only the researcher will know who you are, and that is so that we can send you the survey, or avoid interviewing you twice. We will destroy any reference to you, or your company, in our files after we have sent you the survey.

No ___(265)

Thank you again for your help. You, and your company, will be kept anonymous in any report associated with this interview. Only the researcher will know who you are, and that is so that we can avoid interviewing you twice. We will destroy any reference to you, or your company, as soon as this survey is completed.

This page intentionally left blank.

APPENDIX C

Telephone Log

Telephone Log

No.	Company Name	Contact Person	Telephone Number

Date:_____ Time:_____

No.	Company Name	Contact Person	Telephone Number

Date:_____ Time:_____

No.	Company Name	Contact Person	Telephone Number

Date:_____ Time:_____

APPENDIX D

List of Reasons for Not Collecting Data

Answers to Open Ended Questions.

Reasons Why Training Programs Are Not Evaluated (HQ_{17})

The following reasons were given for not carrying out training evaluations. The listed reasons are paraphrased from the verbal reasons given. The comments were paraphrased because of space limitations in the data base program used to store the research data:

- Programs are too new. Plan to start essay response and on-the-job performance measures.
- We send them to universities...there is no need for evaluation.
- Because we are a very small organization.
- Not to that point yet. It's on the drawing board right now.
- We don't have formal training programs as such. Only special purpose programs.
- We are not mature - we are relatively new to the serious training mentality.
- Until recently no one specifically in charge of training, will do annually.
- Resources, not enough staff or time.
- We do informal follow up, that is adequate, we are a small company.

- We do collect some data at end of training, but impressions only.
- We just hired a training coordinator, and program isn't old enough to evaluation.
- It is done by parent office, or corporate headquarters.
- We go into program with high level of certainty it will be effective.
- We don't know how, and if we did we wouldn't have time or money.
- Only have orientation programs, handle problems as they appear.
- We're very informal, I think it is a mistake, wish we had more formal program.
- So small we don't need formal evaluation program.
- We don't have time to do necessary pre-test, post-test. not evaluated by our department (new franchises).
- Train only basics, no need to evaluate them.
- Training is informal OJT, no need for formal evaluation.
- 1 person for 1000 employees, I get informal feedback.
- Lack of staff.
- Do limited training. Orientation is 1 on 1, & feedback is verbal & informal.
- We only have small orientation programs, we know employees job performance.

- We have orientation and job skill training programs, but no evaluation.
- We are a small personnel department & do not have the manpower to do evaluations.
- We train informally, and effectiveness would be very hard to gauge.
- Just recently been acquired, and don't have the manpower for evaluation.
- We just don't have a need for evaluations at this point.
- We don't have the resources.
- No way of evaluating the types of programs we have, & haven't found necessary.
- We really haven't had any. Just starting up training in a formal mode.

Reasons Why Training Programs Are Not Evaluated for Outputs (HO_{18})

The following reasons were given for not carrying out orientation program evaluations in terms of outputs (on the job trainee behavior):

- It isn't necessary for orientation programs.
- Don't know, we're just not into the evaluation that deeply.
- We don't do it, but in an informal sense.

- On the job effects are not obvious.
- It's not a compulsory program.
- We observe, but don't keep files on this sort of thing.
- We don't know how an orientation program's effects can be measured on the job.
- I get feedback from supervisors, but not in a formal process.
- We don't have a mechanism for collecting that data.
- We run a follow up orientation, but no formal evaluation.
- Once they are trained, that's it.
- Lack of time, difficulty in doing so, lack of interest in data.
- Currently do a subjective, data is not collected numerically.
- Because of limited staff and resources.

The following reasons were given for not carrying out job skill training program evaluations in terms of outputs (on the job trainee behavior):

- I don't know, we just haven't done that as yet.
- Up to supervisor. The training department has no authority to do that.
- Would like to but not enough resources.
- We can't go back into the job and measure some of the job skills.

- Lack of time, difficulty, lack of interest.
- We just haven't gotten to that yet.

The following reasons were given for not carrying out management development program evaluations in terms of outputs (on the job trainee behavior):

- No, only company wide improvements.
- We're not into it that deep.
- Training done by consultant, and evaluation not one of services offered.
- It's usually a time constraint.
- It takes too long to assess that, it's not as quantitative as skills area.
- It's done, but not officially. Hear from managers, but no official data.
- But not for all of the programs.
- We would like to but can't because of limited resources.
- It's real hard to do. We evaluate job performance, but .. not training effect.
- Program is fairly new, & I have sole responsibility for this program.
- Don't have a mechanism for collecting that data.
- We've talked about doing it. Time is a big factor...things too hectic.

- We're working toward doing that, at present time lack tools to do so.
- Once been trained, that's it - don't know why.
- Have no idea, we just haven't done it yet.
- Lack of management interest.
- Just not the policy, observed but not formally evaluated.
- Only subjective evaluations are available.
- We just haven't gotten to that yet.
- Because this would be redundant with another review process we have.
- The manager evaluates the employee after training, but not formally.

The following reasons were given for not carrying out wellness training program evaluations in terms of outputs (on the job trainee behavior):

- The health programs are not specific to on the job behavior.
- We don't have a mechanism for collecting that data.
- ..not like being trained to do a specific job.
- I don't know why, we just haven't.

Reasons Why Training Programs Are Not Evaluated for Outcomes (HO_{19}):

The following reasons were given for not carrying out orientation program evaluations in terms of outcomes (societal impact of learned behaviors):

- We don't know how to measure how orientation effects these outcomes.
- Don't think we have a need in this area .
- Don't relate easily to those outside effects.
- Our department wouldn't do, but believe corporation would.
- Everyone needs electricity. As long as we keep rates down, we're O.K.
- We're distant from our customers, maybe we should have better handle on that.
- We haven't been in existence long enough.
- We are not mature enough in that training area, plus don't have vehicle to do.
- It's not deemed appropriate; too many steps removed.
- We just don't feel that it is relevant.
- We don't have a lot of those training programs.
- I don't know, we just don't do it.
- No demand for data.

- Our training program is new, and not as fully developed as we intend it to be.
- We don't deal with the public, we're a manufacturing company.
- Because of lack of resources.
- The end customer wouldn't have any input.
- Orientation programs are internal.
- Don't know how to do it, and don't have the time.
- There is no way to measure it. We're a manufacturing company.
- Our customer is the government. We don't sell to the public.
- We don't feel it's applicable whatsoever.
- We don't do attitude surveys, unless we can act on results they are damaging.
- I don't know how to do that.
- Would be collected by marketing area if at all.
- I just don't know.
- Because of limited staff and resources.
- Internal is the priority.
- I don't know, except for probably time constraints.
- This, too, is an in-house concern.

The following reasons were given for not carrying out job skill training program evaluations in terms of outcomes (societal impact of learned behavior):

- We never thought about it, or how to go about it.
- We wouldn't do that, but believe the corporation would monitor that.
- Haven't utilized that yet, still in early growth of our training.
- Everyone needs electricity. As long as we keep rates down we're O.K.
- We're not a mature company.
- We don't have a mechanism for collecting that data.
- Bulk of trainees have no contact with customers.
- I don't think it has any impact on our business.
- It's too difficult to measure.
- It's hard to do.
- No demand for data.
- Training programmers don't have time to go back and evaluate this.
- Our program is new, and not as fully developed as we intend it to be.
- We don't deal with the public, we're a manufacturing company.
- We simply don't have the money to do that.
- Because we lack the resources to do this.
- Don't know why.
- The customer would have no input.
- I don't think we need it.

- There is no way to measure it. We're a manufacturing company.
- We don't do attitude surveys, unless we can act on results they are damaging.
- I don't know how to do that.
- It doesn't relate well..job skill training is very internal in nature.
- We just haven't gotten to that point yet.
- Again, too many variables involved.
- Internal is the priority.
- Because we're only concerned with how training effects in house performance.

The following reasons were given for not carrying out management development program evaluations in terms of outcomes (societal impact of learned behaviors):

- We don't know how to measure these affects.
- Basically, its "no" because we get it indirectly in other ways.
- Frankly it never occurred to us.
- I've never really thought about doing and how to go about doing it.
- unless we can act on the results, its more damage than not have done at all
- We don't do, but have to believe the corporation would monitor that.

- In the early growth of training & have not progressed to external evaluation.
- Not old enough to see those effects (2 years).
- Decisions for that information are greater attuned to corporate environment than what we are.
- We haven't developed way to do that yet.
- Too difficult to measure.
- Its hard to do. We do enhance company's personality but don't have data.
- Not sure why, we just don't do it.
- We know the effects without actually collecting data.
- No we haven't gotten that far along.
- Too many uncontrollable variables.
- No demand for that.
- Small company (116 employees.), new trainer, new program, still developing.
- we don't deal with the public, manufacturing only.
- Not enough resources/money.
- We lack resources to do this.
- We have no competition (only hospital in 20 miles).
- We don't know how to do it, and we don't have the time.
- There is no way to measure it, we're a manufacturing company.
- In our business, I don't really know how to do that.
- The marketing area would collect this type of information.

- Too many variables to attribute any particular effect to management training.
- Primarily because of limited staff and other resources.
- Internal evaluation is the priority.
- We don't deal with customers directly.
- I don't know, I guess time constraints.

The following reasons were given for not carrying out wellness training program evaluations in terms of outcomes (societal impact of learned behaviors):

- I don't know how to relate this type of training to client satisfaction.
- Haven't utilized that yet, still in early growth of our training.
- We don't have a mechanism for collecting that data.
- Basically, we're not interested in that information.
- It is hard to know, or get the data.
- We haven't gotten that far.
- We don't deal with the public, we're a manufacturing company.
- Our customer is the government. We have no outside contacts.
- I don't know how to do that.
- Again, this is mainly an in-house concern.

Reasons Why Learner Reaction Data is Not Collected:

The following reasons were given for not carrying out orientation program evaluations in terms of learner reactions:

- Don't know, maybe we should but haven't found the need to do so.
- We just haven't done it, maybe we should.
- If they have any questions, they come see me.
- We're not quantitative, we observe effects on the job.
- A good question, I don't know why we don't, I don't know that it's appropriate.
- We know the instruction is consistent and correct.
- I don't know, we just get verbal feedback.

The following reasons were given for not carrying out job skill training program evaluations in terms of learner reactions:

- We do but very informally, not really a collection of data.
- We just haven't gotten to that yet.
- Don't know specifically, we just get feedback verbally.

The following reason was given for not carrying out

management development program evaluations in terms of learner reactions:

- It's lecture & demonstration, & material is basic & generic. The following reasons were given for not carrying out wellness training program evaluations in terms of learner reactions:
- Again, it was not offered by the consultant.
- We just have never done that.
- We're not looking for that type of feedback - not now anyhow.
- To go back and ask is too personal as far as we're concerned.
- Programs are self paced, and offered as services.
- I don't really have a good answer for you.

Reasons Why Learner Mastery Data is Not Collected:

The following reasons were given for not carrying out orientation program evaluations in terms of learner mastery:

- It isn't necessary for orientation programs.
- We just haven't done it.
- It's not a compulsory program.
- I don't know that it's appropriate.
- We're not at the point where we can capture that.
- We don't have a program to identify that yet.

- We run a follow up orientation, but no formal evaluation.
- I just don't know why.
- I don't really know.
- Lack of time, difficulty in doing so, lack of interest in data.
- Our training program is still too new.
- Newly hired are evaluated on job performance, not on learning the orientation.
- A lack of time.
- Because of limited staff and resources.

The following reasons were given for not carrying out job skill training program evaluations in terms of learner mastery:

- I'm not sure.
- There hasn't been any criticism so we can assume it's doing a good job.
- Will show up in performance appraisal.
- Lack of time, difficulty, lack of interest.
- We evaluate in-house programs, do not evaluate outside vendors.
- Training program still very new.
- No reason given.
- We just haven't gotten to that yet.

The following reasons were given for not carrying out management development program evaluations in terms of learner mastery:

- We evaluate only company wide goals such as sales increase.
- I don't know, we just don't.
- Training done by consultant, and evaluation not one of services offered.
- We don't have time or money.
- It'll show up in performance appraisal, so employees aren't tested.
- Have no idea, we just haven't done it yet.
- We haven't determined viable way of doing that yet.
- difficulty, time constraints, lack of management interest.
- Just not done, that's all I can say.
- We lack resources to do this.
- Because material is basic and generic rather than a formal degree or program.
- Because corporate headquarters collects the data.
- We just haven't gotten to that yet.
- We just haven't taken the time for this kind of follow up.
- We haven't felt the need to follow that up.

The following reasons were given for not carrying out

wellness training program evaluations in terms of learner mastery:

- It was not offered by the consultant.
- We feel it's unnecessary.
- We don't feel it is necessary.
- I don't know why, we just haven't.

BIBLIOGRAPHY

BIBLIOGRAPHY

Baker, R.L. and Schutz, R.E. (Eds.). (1972). Instructional Product Research. Cincinnati, Ohio: Van Nostran Reinhold Co.

Bertalanffy, L. (1968). General System Theory. New York: George Braziller Publishers.

Bertalanffy, L. (1962). General System Theory: A Critical Review. General Systems, VII., 1-20.

Brethower, D.M. (1966). A Systems Approach to Educational Problems. Occasional Paper, Center for Programmed Learning for Business, University of Michigan.

Brethower, K.S. & Rummler, G.A. (1979, May). Evaluating Training. Training and Development Journal.

Cohen, A. & Hyman, J.S., (1979, December). How Come So Many Hypotheses in Educational Research are Supported? Educational Research, 12-16.

Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences. (Rev. Ed.) New York: Academic Press.

Corrigan, G.J. (1980, January). Corporate Training: A Career for Teachers? Phi Delta Kappan, 61(5), 328-331.

Cunningham, D.J. (1971, February) Formative Evaluation of Replicable Forms of Instruction. Paper presented at the annual meeting of the American Educational Research Association, New York. Fitz-Enz, J. (1984). How to Measure Human Resources Management. N.Y.: McGraw-Hill, Inc.

Fitz-Enz, J. (1985). Human Resource Effectiveness Survey: 1st Quarter 1985. Saratoga, Calif.: Saratoga Institute. Fowler, F.J. Jr. (1984). Survey Research Methods. Beverly Hills, California: Sage Publications, Inc.

Gael, S. (1983). Job Analysis. San Francisco, Calif.: Josey-Bass. Gilbert, T.F. (1978). Human Competence: Engineering Worthy Performance. New York: McGraw-Hill.

Hays, W.L. (1981). Statistics. New York: Holt-Rinehart & Winston.

Horrigan, J.T. (1979, July). The Effects of Training on Turnover: A
Cost Justification Model. Training and Development Journal, 3-8.

Kaufman, R. and Stone, B. (1983). Planning for Organizational Success. New York: John Wiley & Sons.

Kaufman, R. and Mayer, H. (1981, February). Educating for Beyond Performance and Instruction: A Pragmatic Appraisal. NSPI Journal, 8-12.

Kaufman, R., Mayer, H., and Butz, R. (1984, May). Defining and Classifying the Organizational Elements. Performance and Instruction Journal, 22-25.

Kearsley, G. (1982). Costs, Benefits, & Productivity in Training Systems. Reading, Mass.: Addison-Wesley.

Latham, G.P. & Wexley, K.N. (1982). Increasing Productivity Through Performance Appraisal. Reading, Mass.: Addison-Wesley.

Lumsdaine, A.A. (1965). Assessing the Effectiveness of Instructional Programs. in Glaser (Ed.) Teaching Machines and Programmed Learning. Washington, D.C.: NEA.

Lusterman, S. (1977). Education in Industry. New York: The Conference Board.

Mager, R.F. (1984). Measuring Instructional Results. Belmont, California: Pitman Learning, Inc.

Mager, R.F. & Pipe, P. (1984). Analyzing Performance Problems. (2nd Edition) Belmont, California: Pitman Learning, Inc.

Markle, S.M. (1967). Empirical Testing of Programs. Sixty-Sixth Yearbook of the National Society for the Study of Education, Part II, 104-138.

Nadler, Leonard. (1979). Developing Human Resources, 2nd ed. Austin: Learning Concepts.

Nie, N.H., Hull, G.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. (1975). Statistical Package for the Social Sciences, 2nd Ed.

Popham, W.J. (1973). Evaluating Instruction. Englewood Cliffs, N.J.: Prentice-Hall.

Posavac, E.J. and Carey, R.G. (1980). Program Evaluation: Methods and Case Studies. Englewood Cliffs, N.J.: Prentice-Hall.

Rogers, E.M. and Shoemaker, F.F. (1971). Communication of Innovations: A Cross-Cultural Approach. N.Y.: Free Press.

Rummler, G.A., (1985). Organizations as Systems: Introduction and Basic Premise. Summit, New Jersey: The Rummler Group.

Sanders, J.R. and Cunningham, D.J. (1972). A Structure for Formative Evaluation in Product Development. Educational Research and Evaluation Laboratory, Indiana University.

Scriven, M. (1967). The Methodology of Evaluation. AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally and Co., 39-83.

Shipp, T. (1980, November). To Survive the Budget Inquisition, Prove Your Training Makes Dollars and Sense. Training/HRD, 23-29.

Sibson, R.E. (1985). The Management of Personnel. N.Y.: Vantage Press.

Stake, R.E. (1970). Objectives, Priorities, and Other Judgment Data. Review of Educational Research, 40, 181-212.

Thompson, M.S. (1980). Benefit-Cost Analysis for Program Evaluation. Beverly Hills, Calif.: Sage.

Tyler, R.W. (1931). What is Statistical Significance? Educational Research Bulletin, 10, 115-118.

Wilkensen, G.L. (1972). Cost Evaluation of Instructional Strategies. AVCR, 21.

Worthen, B.R. and Sanders, J.R. (1973). Educational Evaluation: Theory and Practice. Worthington, Ohio: Charles A. Jones Publishing Co.