6-1993

The Importance of Evaluation Questions in Human Resources Development: Audience Perspectives

Fahad Al-Ruwaished

Western Michigan University

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

Part of the Educational Assessment, Evaluation, and Research Commons, and the Educational Leadership Commons

Recommended Citation

http://scholarworks.wmich.edu/dissertations/1883

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE IMPORTANCE OF EVALUATION QUESTIONS
IN HUMAN RESOURCES DEVELOPMENT:
AUDIENCE PERSPECTIVES

by
Fahad Al-Ruwaished

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

Western Michigan University
Kalamazoo, Michigan
June 1993
The purpose of this study was to investigate and explore how the perception and focus of human resources development (HRD) evaluation relates to different audiences and clients, and to make recommendations to individuals and organizations involved in designing, developing, and implementing HRD evaluation. The study used the CIPP (Context, Input, Process, and Product) evaluation model to operationally define the kinds of evaluation questions as dependent variables. The study independent variable was the organizational role, operationally defined as simulated HRD audiences, i.e., CEOs, managers, trainers, and trainees.

The human resources evaluation questionnaire (HREQ) was developed by the researcher to measure the perceived importance of evaluation questions. The questionnaire went through three stages of development resulting in a reliable and valid measure of the dependent variables. The study sample consisted of seventy-three graduate level students attending the Human Resources Development program at Western Michigan University.
The findings of the study were based on the testing of four hypotheses. The hypotheses were tested by ANOVA one-way analysis of variance. The results of the data analysis provided no evidence that the role of HRD audiences influences the perceived importance of evaluation questions. The four dependent measures were substantially correlated.

Based on the findings of the study, it was concluded that there appears to be a high relationship among the four types of categories in the CIPP evaluation model. The HRD practitioners perceived all information to be important for their evaluation. The HREQ instrument can be utilized as planning tool for HRD evaluation.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The importance of evaluation questions in human resources development: Audience perspectives

Al-Ruwaished, Fahad A., Ed.D.

Western Michigan University, 1993
ACKNOWLEDGMENTS

There are a number of people who deserve special recognition for their interest and assistance in the preparation of this dissertation.

A special expression of gratitude is due Dr. Mary Anne Bunda who served as chairperson of my doctoral committee. Her expertise, guidance, and support not only in the preparation of this dissertation, but throughout the doctoral program made me realize that I can reach my goals only through dedication and hardwork. My degree program has been interrupted by the second Gulf War which caused me a delay of two years and got me close to giving up pursuing my degree. If there was one person that I can say was behind my coming back to pursue my degree, it will be Dr. Mary Anne Bunda.

I would also like to thank Dr. Eugene Thompson and Dr. Alyce Dickinson for their helpful input as committee members in providing numerous suggestions for the improvement of this manuscript.

A special expression of love and appreciation goes to my beloved mother Nora Borhama, and my father Abdulrahman Al-Ruwaished, who both believed so strongly and given me so unselfishly. A special expression of love to my brothers
Acknowledgements—Continued

Mohammed, Naji, Salaman, Abdulla, Addle, and my lovely sister Kholood.

Finally, I wish to pay tribute to dear friends Baqer Dashti, Mohammed Qadiri, Sultan Yaquob, and Fawzi Al Majdali who all have consistently and cheerfully supported me throughout the doctoral program.

Fahad Al-Ruwaished
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** .................................................. ii  
**LIST OF TABLES** .................................................... vi  
**LIST OF FIGURES** .................................................. vii  

**CHAPTER**

I. **INTRODUCTION** .................................................. 1  
   Human Resources Development .......................... 4  
   Evaluation Within the HRD Context .................. 5  
   Purpose of the Study ................................. 8  

II. **REVIEW OF RELATED LITERATURE** ............................. 9  
    Evaluation of Theoretical Literature ............ 10  
    Audience Identification ........................... 15  
    Models and Goals of HRD Evaluation ............... 18  
    Summary of the Chapter ........................... 27  

III. **METHODOLOGY** .................................................. 30  
    Identification and Description of Variables .... 30  
    Instrument Development .......................... 31  
    Instrument Plan ........................................... 32  
    Panel of Experts: A Validity Study ............. 33  
    Pilot Test: An Estimate of Reliability ........... 35  
    Simulation Development .......................... 39  
    Design and Data Collection ....................... 41  
    Operational Procedures ........................... 42  

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

CHAPTER

Hypotheses ...................................... 43
Data Analysis Procedure ........................ 44

IV. RESULTS OF THE STUDY .......................... 46
Instrument Characteristics ...................... 47
Test of Hypotheses ............................... 49
Correlational Analysis .......................... 53
Summary of the Chapter ........................... 55

V. CONCLUSIONS AND RECOMMENDATIONS ........ 56
Discussion of Findings .......................... 56
Limitations of the Study ......................... 58
Recommendations for Future Research .......... 61
Conclusion ......................................... 63

APPENDICES ............................................. 65
A. Letter of Request to the Content Validity Panel... 66
B. Pilot Survey Instrument ........................ 74
C. Survey Instrument ............................... 83
D. Simulation Role for Participants .................. 89
E. Request for Subject Solicitation and Script ..... 94
F. Human Subjects Institutional Review Board .... 97

BIBLIOGRAPHY ............................................ 99
LIST OF TABLES

1. Classification of Items in the Initial Draft ........ 33
2. Reliability Coefficient for the Pilot Test .......... 38
3. Change in Instrument After Item Review .......... 39
4. Characteristics of the Dependent Variables Measures ................................................. 48
5. Analysis of Variance Summary for the Context Category .............................................. 50
6. Analysis of Variance Summary for the Input Category ................................................... 51
7. Analysis of Variance Summary for the Process Category .................................................. 52
8. Analysis of Variance Summary for the Product Category .................................................. 53
9. Correlation Coefficients for the Four Types of Information ............................................. 54
LIST OF FIGURES

1. Six Stage Model by Brinkerhoff ...................... 25
CHAPTER I

INTRODUCTION

Evaluation is a systematic process to determine the worth, value, or meaning of programs (Joint Committee, 1981). Evaluation refers to the activity of systematically collecting, analyzing, and reporting information to specific identifiable interest groups for the purpose of making decisions. This definition focuses on gathering information that is meant to be used for program improvement and decision making (Alkin, 1990). The definition indicated systematic which means that the evaluation must be planned. This plan should be aimed at obtaining information that will answer specific questions of a specified potential audience.

Evaluation of programs is conducted for many purposes or decisions. Evaluators often design either a formative and/or summative evaluation system. Formative evaluation is defined as ongoing evaluation that focuses on the implementation of the program. It improves the effectiveness of a specific program while it is in the implementation stage. Program directors can modify program elements on the basis of this feedback. Summative evaluation on the other hand, is defined as outcome evaluation designed to improve future
but similar programs. The purpose of summative evaluation is to describe the flow of activities, and the achievement of goals and objectives (Scriven, 1967). Thus, the audience for formative evaluation is primarily internal, while the potential audiences for summative evaluation are primarily external to the program.

In addition to the formative/summative distinction, several models have been generated to describe the questions that focus on evaluation. The CIPP evaluation model was designed to provide both formative and summative data as elaborated by Stufflebeam in the late 1960s and will be the basis for discussion in this study. According to the CIPP model, there are four categories of foci to consider in evaluation: Context, Input, Process, and Product. It looks at evaluation as the process of planning, obtaining and providing useful information for judging decision alternatives at four distinct phases of program development (Stufflebeam, 1974).

The entire focus of the CIPP model is improving the effectiveness of educational programs by providing a framework to analyze the type of information needed in relation to the phase of program implementation (Galvin, 1983). Context evaluation is concerned with identifying the strengths and weaknesses of a program to provide information about needs, problems and opportunities in order to identify and judge objectives. Input evaluation is in-
tended to look at program strategies, and provide information for an existing program in order to assess whether or not the program will work for the organization's benefit. Process evaluation focuses on the implementation of the program to provide feedback about the program activities. Product evaluation deals with attainment of objectives. It provides information for determining whether objectives are being achieved and whether the procedures employed to achieve them should be continued, modified or terminated (Galvin, 1983; Stufflebeam & Shinkfield, 1985).

Although theoretical evaluation models have been formulated to assure comprehensive designs, in practice programs are rarely evaluated in such a way that the data provide adequate information on the effectiveness of programs or the ways they can be improved. Evaluators are often told that their evaluation can not be used because of bureaucratic relationships and conflict. The evaluation can only be used if management really wants the data. Evaluation often seems to be too late to be useful, too full of jargon to be understood, and too lengthy for the user reading time available. Therefore, evaluators need to learn that they must be useful to others. They must understand the political system in which evaluation occurs, and the information needs of decision makers who utilize evaluation (Phillips, 1983). Quality evaluation must take into consideration not only the phase of program implementation,
but also decision-maker needs.

Human Resources Development

Human resources development (HRD) is a topic frequently discussed in training and development literature. Many prescriptive models have been developed for conducting HRD evaluation, but little is known about the effectiveness of HRD programs. HRD as Nadler (1980) explains, can be constructed as "Training" when its primary purpose is to improve current job performance, "Education" when its primary purpose is to help personnel advance to different jobs, and "Development" when it aims to strengthen the organization through benefiting the individuals.

The bottom line of almost all HRD programs according to Brinkerhoff (1987) is the organizational benefits, in other words, the impact of HRD on the organization. Organizational benefits can be measured in a variety of forms: high morale, low turnover, more commitment to the organization, and above all an increase in productivity. It is important to understand that some HRD programs benefit the organization through on the job application of skills, knowledge, and attitude, but this is not always the aim of the HRD program (Brinkerhoff, 1987). The following two examples illustrate this difference. An HRD program can aim to improve the efficiency of the sales force over a given period of time. Once the sales people learn the new
skills, knowledge, and attitudes required to make these sales, the benefit to the organization is an increase in productivity (Hamblin, 1974). The other example is a retirement training program to increase skills and knowledge in retirement planning. This type of training might not benefit the organization directly however, it shows the individuals that the company cares, which ultimately leads to more commitment and loyalty towards the organization (Brinkerhoff, 1987). Consequently, HRD programs have a variety of criteria that can be applied in evaluations.

Evaluation Within the HRD Context

The success of many human resource programs has been measured in terms of the number of participants involved not with respect to the stated goals and objectives of the program. Human resource developers generally agree that the problem of accurate and effective evaluation is both complex and far from a satisfactory solution (Fast, 1975). The question is not whether evaluation of training is possible or impossible, but how systematic and comprehensive evaluation can be carried out and by whom and for whom.

While evaluation is important for many interest groups, most people think of evaluating HRD as examining the effects training has had on people and their performance (Camp, Blanchard, & Huszczo, 1986). HRD evaluation literature does not deal with the issue of audience, but
uses other foci of evaluation. Most HRD evaluations are seen as a dimension of Kirkpatrick's evaluation model. According to Kirkpatrick (1975), there are four levels to consider in evaluating training; reaction, learning, behavior, and result (Kirkpatrick, 1975).

The design of an evaluation as suggested by Worthen and Sanders (1987) should begin with the identification of all parties who will be affected by the evaluation data. Evaluation always involves multiple and diverse audiences, therefore, an evaluation planned and conducted to meet the information needs of identified audiences is more likely to be used. Identification of audiences and clients would help evaluators think broadly of the purpose that might be served in providing them with the evaluation information, and to determine what information each audience needs and will use (Brinkerhoff, 1987).

Most evaluations are of concern to multiple audiences, and each audience can be expected to raise a number of different issues concerning the object to be evaluated. Standards of evaluation practice suggest that the scope and selection of information should address questions about the object of the evaluation, and be responsive to the needs and interests of specified audiences (Joint Committee, 1981).

No one evaluation should attempt to evaluate everything about the HRD programs. Rather, Worthen and Sanders
(1987) suggest it should evaluate those aspects of the program that would yield the information required to meet the audiences' needs and purposes. For example, if the evaluation purpose is to determine what parts of the training need revision, participants' reaction and learning should be assessed. But if the purpose of the evaluation is to decide whether a training program is worth keeping, results and outcomes to the organization should be assessed.

The evaluation literature indicates that there are potentially three basic interest groups for any evaluation. Those groups are the clients, the stakeholders, and the audience. The client is the specific agency or individual who requests the evaluation. Stakeholders are those who may be affected by the evaluation results. Audiences are those individuals, groups, and agencies who have interest in the evaluation and receive its results. Sponsors and clients are usually the primary audiences (Worthen & Sanders, 1987). In HRD, the audience may be the CEO of the organization, the HRD manager, or the HRD trainer. Trainees would always be stakeholders in an HRD evaluation. The possibility of different audience groups would suggest that the evaluation may have different foci. These groups may have systematically different interests in the HRD programs, or as Brinkerhoff suggests they may all have a common "bottom line."

Audience and clients have not been addressed in the
HRD literature. The objective of the study is to investigate whether the type of evaluation questions and information regarding the HRD effort have different importance to audiences and clients depending upon their role in the organization.

Purpose of the Study

The purpose of evaluation and the identification of audiences should be one of the most important elements of the whole process of HRD evaluation. However, most HRD evaluation literature focuses on reaction and learning of participants suggesting a common value for program success, whereas evaluation theory and standards focus more and more on audience and clients and how they perceive the information needed. In evaluating human resource development programs, the audiences can be the person who delivers the training, HRD managers, and CEO of the company. But it is unknown whether these audiences have systematically different perspectives on the value of information for decision making.

This study investigated whether the priorities of the focusing questions of an evaluation depend upon audience role in the organization. More specifically, the purpose of this study was to investigate and explore how the perception and focus of HRD evaluation relates to different audiences and clients.
CHAPTER II

REVIEW OF RELATED LITERATURE

The development of evaluation models has expanded rapidly over the past three decades. More models and strategies for improving the effectiveness of evaluation have been and continue to be developed. The evaluation literature can provide a number of useful concepts for the articulation of the question "How does one do evaluation?" First, what is evaluation, second why one does evaluation, and third how does one make judgments. Some of the terminology developed by evaluation has been accepted by a wide range of authors and is common to all applied evaluation literature. Other languages have been developed specifically by proponents of one system of evaluation or authors. The common languages set out broad issues of agreement and disagreement.

Three dichotomies of concepts over-ride all unique evaluation models and speak to the broad question of how evaluation is completed. These issues are formative/summative or what sort of questions should be asked, merit/worth or how judgments should be made, and impact/product or what focus the evaluation should have. The last issue represents the one that generates the most dis-
agreement in the field possibly because it is the vaguest.

This chapter clarifies the distinction between the evaluation literature and evaluation within the HRD context. It opens with an introduction of the different types of evaluation: formative and summative evaluations, responsive evaluation, the CIPP evaluation model; the use of terms merit and worth in evaluation; and evaluation impact. Audience identification, models and goals of HRD evaluation are also discussed.

Evaluation of Theoretical Literature

Scriven (1967) introduced the terms formative and summative evaluation which have received considerable attention in the research and literature since their introduction. Formative evaluation is the process to obtain data in order to revise programs to make them more efficient and effective. Such activities are oriented toward feedback and developmental assistance which occur prior to full field implementation of programs.

Summative evaluation refers to the activities that are conducted after the field implementation of programs. The purpose is to determine the effectiveness, efficiency, and value of the programs. The activities of summative evaluation are oriented toward the impact of programs, such as the impact of training, cost savings, and return on investment (Scriven, 1967). This early distinction between the
formative and summative evaluation provided evaluators with a mechanism for selecting appropriate criteria for data collection. This separation provided the evaluators with why one does evaluation, who will use the results, and how they will use those results. This distinction concerns the major use of the information, but does not provide a basis for selecting specific focusing questions for the evaluation.

Stake (1975) developed and presented responsive evaluation which is less concerned with the objectives of the program. Stake (1975) prefers to view evaluation as a way to perform a service that can be useful to specific audiences. Responsive evaluation suggests that in order for the evaluation to be useful, evaluators should know the interests and the language of their audiences. The evaluator should have a good sense of for whom he/she is working and the information needs of the audiences for whom the evaluation is being done. It provides information that is relevant to the audience needs and utilization.

Evaluation can serve many different purposes, for example, to document events (the summative role), or to aid in decision making and facilitate remediation (the formative role). The purpose of any given evaluation should be determined by the different information needs of different audiences. There are many steps in conducting responsive evaluation. The evaluator begins with the identification
of audiences and program staff to gain a sense of their posture with respect to the purpose of evaluation. As a result, the evaluator places limits on the scope of the program that need to be evaluated, and begins to conceptualize the problems that the evaluation should address (Stake, 1975).

Stufflebeam (1974), in the 1960s, developed the CIPP evaluation model. As explained earlier, there are four levels of evaluation within this model: Context, Input, Process, and Product. Context evaluation is intended to identify the strengths and weaknesses of the programs' needs and objectives. Input evaluation is intended to look at program strategies, and provide information for an existing program in order to assess whether or not the program will work for the organization's benefit. Reviewing the practices, procedures, and schedules of the program with respect to meeting the specified objectives is the method for this evaluation. Process evaluation focuses on the implementation of the program to provide feedback about the program activities. The methods for process evaluation are also varied and depend on the objectives. The evaluator might review the program materials, and observe the activity while the program is in the implementation stage. Product evaluation deals with attainment of objectives. It provides information for determining whether objectives are being achieved and whether the procedures employed to
achieve them should be continued, modified or terminated. Methods for product evaluation consist of surveying a selected sample of the participants, asking how the program has influenced their work, and observing the performance of participants. In general, context evaluation assists in forming objectives, input evaluation aids in program planning, process evaluation guides implementation, and product evaluation helps in recycling decisions (Galvin, 1983; Stufflebeam & Shinkfield, 1985). While different groups may be involved, the focus of the CIPP evaluation model is on internal program development and it results, rather than the interest of parties.

Unlike other models, the CIPP evaluation model is a future oriented model for decision-making activities (Isaac & Michael, 1981). It is also worth mentioning, that the CIPP model has some weaknesses. For example, evaluation does not always provide all information upon which to base a decision; evaluation for the sake of decision making might miss the opportunity for greater impact; and evaluation results or information are always subject to varying interpretations (Brinkerhoff, 1981).

Guba and Lincoln (1981), using an initial distinction provided by Scriven (1967), argued that evaluation involves both describing and judging activities to place a value on programs being evaluated. Value can be described as merit or worth. They both mean value, but there is a distinction
between them. For clarification, consider gold as an example to differentiate between the two terms. Gold might be judged for merit on its beauty that permits it to be fashioned into such things as jewelry, whereas it might be judged for worth in the trading market.

Evaluation of merit is made in terms of criteria that are relatively stable over time. These criteria come from the literature of a field. Evaluation of worth is made with criteria that may alter rapidly with changing conditions. These criteria in contrast to merit criteria come from individual audience or context perspectives. One strategy of evaluation is to determine the merit of a program in a number of different situations and then to describe its worth in each of those different contexts.

According to Guba and Lincoln (1981), there are two ways to determine the merit of programs. First, by determining the degree to which it conforms to certain standards as dictated by outside agencies. Second, by comparing the program with other programs within the same class or level. The first is called absolute merit evaluation, and the second is called comparative evaluation. Worth, on the other hand, is determined by comparing the programs' outcomes to selected sets of external requirements (Guba & Lincoln, 1981). The importance of the merit/worth distinction is that it clearly shows that different values may be ascribed to an object.
Evaluation impact has been used differently by many authors within the fields of applied evaluation and lacks the common use and precision of the merit and worth dichotomy. Rossi and Freeman (1989) discussed evaluation impact as an effort to establish causality to determine whether the proposed program has its intended effects. Impact evaluations might be used to test the most effective ways to develop and integrate the various program elements.

Impact evaluation compares targeted measures that result from an intervention with the same measures from a control or nonintervention situation. Comparisons might be examined between measures of the behaviors of individuals and conditions, or between measures of outcomes that would have occurred in the absence of intervention (Rossi & Freeman, 1989). It is important to note that impact evaluation usually concerns broad goals not specific objectives of a program.

Audience Identification

Evaluation models are differentiated on the basis of their primary organizers. However, all models include some discussion of audience or clients. Posavac and Carey (1992) view evaluation as a goal based effort for determining whether the program meets the needs of clients and stakeholders. Evaluation is perceived as a tool to help
clients find discrepancies between the goals and the needs of the target population. An effective evaluation plan usually begins with the identification of clients and stakeholders. Clients and stakeholders are defined as those who are personally involved in the program or affected by the quality of the program.

In planning an evaluation, program participants should be identified as one of the first sources of information. Participants who receive a service are often in a good position to provide information about its many aspects. They are the most knowledgeable people about their current status, and can provide self assessments for a variety of dimensions.

Program staff are the second source of information. They are trained to assess participants' improvement and how well programs are being managed, and thus they can provide good data for program evaluation. Therefore, evaluators need to recognize that program evaluation is an evaluation of participant improvement and staff performance (Posavac & Carey, 1992). Evaluation of staff performance is not the same thing as having the staff evaluate the participants or the program. Evaluation of staff performance would consist of the supervisor of the staff evaluating how well they implemented the program.

Nevo (1974) conducted a study that focused on how education audiences perceived the importance of evaluation,
and what type of information is available to them. One of the purposes of the study focused on identifying the audiences of evaluation in the school system (principals, teachers, and students), and what type of information they perceived as the most appropriate to be considered for the development of an evaluation system using the CIPP and formative/summative evaluation models. Based on the findings of the study, the scope of the evaluation and how audiences perceived evaluation were the two criteria suggested to be used in choosing an evaluation system. In the data analysis there was a significant interaction between the availability of information for the data obtained from students and principals. No interaction effect was found in the teacher group. With regard to the differential importance of information, the results of the study suggested that there was a difference between the relative importance students and principals attached to formative/summative evaluation. No difference was found for the teacher group.

The evaluators are always focusing on clients and audience needs and concerns. Some audiences want to see information related to the achievement of objectives, another audience wishes to influence service decisions. Therefore, an evaluation should begin with audience identification and the information needs of those audiences.

"Audience involved in or affected by the evaluation should be identified, so that their needs can be addressed"
In the list of professional criteria to be used to judge evaluation, audience identification is the first standard for evaluation, consequently the most important one. Most evaluations are of concern to multiple audiences, and each is expected to raise a number of different issues concerning the object to be evaluated. Since some audiences will usually be more important than others, some weighting of their input may be necessary. Identification of audience and clients helps evaluators think broadly of the purpose that might be served in providing them with the evaluation information, and to determine what type of information each audience needs and will use (Worthen & Sanders, 1987).

It is important to note that the choice of evaluation data sources depends on the type of decision to be made on the basis of that information. Therefore, evaluators should strive to use information from multiple sources in order to have a valid and useful evaluation.

Models and Goals of HRD Evaluation

Government agencies, industrial firms, and educational institutions are placing more and more emphasis on the need to develop their human resources. This pressure is enlarging the role and responsibility of training and development professionals, particularly with respect to providing experiences, processes, and programs that will
lead to individual and organizational development (Lippitt, 1975). Too often HRD programs have not followed logical steps in their design and development. They are solutions looking for problems (Phillips, 1983). In the last few years more and more emphasis is being placed on evaluating the results of HRD programs. The stated goals of evaluation are to improve the HRD programs, to provide feedback to program planners and instructors, and to assess the trainees skills, knowledge and attitudes (Swierczek & Carmichael, 1985).

However, evaluation of HRD may not be meeting these broad goals. The literature indicates that there are only a few reports of successful HRD evaluations. Several authors have categorized and defined the large number of HRD evaluation models and procedures, but these models do not specifically include the diverse information needs for multiple audiences, and the focus and scope of these audiences.

Audience identification for human resources evaluation requires careful consideration. Two primary questions should be asked: Who must be satisfied with the evaluation; and on whom will the evaluation have impact? Answers to both questions will provide an initial list of clients and audiences. Each audience will have its own needs relative to the information desired. Some want information to make operational decisions, others need just to be informed.
Phillips (1983) suggests that the most important audiences are the HRD managers and staff. They must receive information about program results. The program designer and developer must have the information on the program's effectiveness. This is necessary so that adjustments can be made for operational decisions if the program is to be repeated in the future (Phillips, 1983).

Trainees need feedback on the overall success of the HRD effort. Funding agencies and sponsors are also important because they are responsible for the allocation of resources; they need information to help justify the expenditures. Professional organizations and other stakeholders need to be informed about the HRD programs results in a more general way so they can build respect for the HRD activities in the eyes of employees (Phillips, 1983).

Audience and clients can play an important role in establishing the evaluation purposes, presenting issues that need to be addressed, and identifying decisions that are to be made. They can help influence specific evaluation objectives, design, and strategies. For instance, without such information, an evaluator may conduct a summative evaluation when perhaps the client really needs a formative evaluation.

The process of identifying the audience and clients that should receive various aspects of information should occur at the planning stage of HRD evaluation. This pro-
cess can be done by relating the evaluation questions to be addressed to the potential users of the information. The following section will present two important HRD models that have been discussed in the general evaluation literature; one developed by Kirkpatrick (1975), and one by Brinkerhoff (1987).

Kirkpatrick (1975) developed an evaluation model that is based on three major assumptions. First, there are four levels for HRD evaluation: reaction, learning, behavior and results. These levels have been accepted as standards by many practitioners with minor modification. Second, the four levels are arranged in the order of the value of information to be gained through evaluation. Third, there is a causal correlation between the four levels of evaluation. If trainees react favorably, they will probably learn more, if they learn more, they will probably change their behavior and if they change their behavior, the usual indices of performance will improve (Newstrom, 1978).

Level I: Reaction can be defined as how well the participant liked the program. It is important to determine how people feel about the program because individuals are motivated in different ways. Trainees who liked the HRD program are more likely to benefit from it. Analyzing participants' perceptions of and reactions to the program in terms of content, presentation and methods of instruction can suggest ways of making improvement to the program. It
can also provide valuable information for improving future programs.

A common and effective way to evaluate participant reaction is the reaction sheet. The human resource developer can design a reaction sheet in a tabulated form which can allow open-ended comments to be written for additional reaction (Kirkpatrick, 1975). The second way involves gathering data on the emotional acceptance of the material taught using a rating scale (Braun, 1979).

Level II: Learning can be defined as principles, facts and skills which are understood and absorbed by the participants of the HRD program. It is generally agreed that the primary value of any HRD program is gaining new skills of job operation and attaining new knowledge to improve the participants' effectiveness. The purpose of evaluation in terms of learning is to determine what changes in skills, knowledge and attitudes occurred as a result of participation in the HRD program (Kirkpatrick, 1975). Successful HRD evaluation in terms of learning is often measured by a pre-post examination. Participation in and outcomes of in-class training can serve as a substitute for the pre-post course examination (Newstrom, 1978).

Level III: Behavior evaluation can be defined as changes in job behavior as a result of the HRD program (Kirkpatrick, 1975). The objectives of most HRD programs are to change behavior in such a way that observable behav-
ior will occur outside the learning environment. Evaluation in terms of behavior is an extremely difficult matter to deal with. This is attributed to the lack of meaningful criteria for judging on the job application. This type of evaluation is usually achieved by defining the performance objectives of the HRD program. Participants should understand that the task of the HRD program is to improve their job performance (Couch & Strohler, 1971).

In order to accomplish the behavior objective, the participants must commit themselves to specific objectives on how they will apply what they have learned in the program and to send a progress report within a specified time after the program. These objectives must be related to the material in the program and there must be a sufficient amount of time for the application of the objectives (Morrisey & Wellstead, 1980).

Level IV: Results can be stated in terms of impact such as higher productivity, changes in attitude, and increases in interest and loyalty (Kirkpatrick, 1975). There are a number of benefits in measuring the results of the HRD program: (a) to justify if the program helps to change behavior, (b) to judge if the HRD program produces impacts on the organization, and (c) to offer the human resource developers an opportunity to demonstrate to management the importance of the HRD intervention in relation to the financial health of the organization (Trapnell, 1984;
Unfortunately, measurement of results and impact is a very difficult task. Many factors can complicate the process in evaluating the results of the HRD program. Measuring results requires considerable front-end analysis to determine if desired objectives such as changes in attitudes, reduction in grievances or increasing loyalty were achieved (Brandenberg, 1982). Therefore, as Kirkpatrick (1975) and others have indicated, it is recommended that human resource developers evaluate in terms of reaction, learning and behavior.

Undoubtedly, the four preceding aspects of HRD evaluation are important, but there is either a misuse or redefinition of evaluation terms such as evaluation impact. Further, because the four measures occur after the HRD intervention, there is also no explicit use for improvement of programs (formative evaluation).

Brinkerhoff (1987), expanded Kirkpatrick's (1975) four levels of evaluation with the development of a Six Stage HRD evaluation model. In evaluating HRD programs one of the key elements of the model is to focus on merit and worth. According to Brinkerhoff merit is the determination of how well programs are done. It is the determination of program impact on the organization, which is substantially different from standard use of the terms. Merit is represented by stages 2, 3 and 4, while worth, on the other
hand, is represented by stages 1, 5 and 6. The process of conducting a comprehensive HRD evaluation is outlined in Figure 1 which depicts each one of the six stages beginning at needs assessment.

![Diagram of Six Stage Model](image)

**Figure 1. Six Stage Model by Brinkerhoff (1987).**

Stage 1 is needs assessment. The evaluation at this stage is to look at the importance of the problem, and how needs can make a difference to the individual and to the organization as well. Needs assessment is the process to facilitate evaluation. The datum required to evaluate this stage is specification of program objectives. It is impossible to conduct a proper evaluation of results without paying careful attention to the objectives of the program.

Stage 2 is program design. HRD programs are designed to meet certain goals and objectives. The evaluator should seek data regarding the input, process, and output. The outputs of the HRD program will enable the evaluator to
measure changes in personal characteristics in order to recognize the impact of the HRD.

Stage 3 is program implementation. It assesses the HRD program while it is in operation. The evaluator should seek data regarding the activities and achievement of the enabling objectives of the program. The data should also highlight the strengths and weakness of the program while it is in operation (formative role).

Stage 4 is concerned with the immediate outcomes or learning. In conducting a HRD program, the underlying assumption is that the participants in the program will learn specific skills, knowledge, and attitudes, otherwise, there would be no point in investing the time, money and energy to design and implement a program. Data required for this stage were the assessment of skills, knowledge and attitudes.

Stage 5 is on the job application. An evaluation of this stage is the observation of how much of the skills, knowledge and attitudes are being applied to the job situation. Data required for this stage should be dealing with changes in behavior.

Stage 6 is organizational benefits. This stage is directly related to Stage 1 needs and objectives. There is a definite linkage between HRD objectives and results. Program results can be stated in terms of how well the training objectives were achieved (Brinkerhoff, 1987; Kohn
& Parker, 1975). The evaluator should seek data regarding the impact of the HRD program on the organization and also, the financial costs and benefits of the HRD program. This type of data are applied to determine whether the HRD program makes economic sense (Brinkerhoff, 1987; Kohn & Parker, 1975; Phillips, 1983).

Brinkerhoff expanded Kirkpatrick's (1975) dimensions to include stages of formative evaluation for instructional design. Thus, he has accurately used one of the standard concepts of evaluation. With regard to Kirkpatrick's model (1975), Stage 4 immediate outcomes relate to Level II learning; Stage 5 on the job application is the same as Level III behavior; and Stage 6 organizational benefits is Level IV results. In relation to the CIPP model, Brinkerhoff's Stage 1 needs assessment falls under Context evaluation; Stage 2 program design falls under Input; Stage 3 falls under Process; and Stages 4, 5, and 6 all fall under Product evaluation. But neither Brinkerhoff (1987) nor Kirkpatrick (1975) explicitly show the focusing capabilities of different audiences for an evaluation.

Summary of the Chapter

Evaluation is not an easy process. The evaluator must design an evaluation system that puts the data in proper perspective. If the evaluation is to be used solely for rewarding the instructors' or the trainees' performance,
the evaluation plan can be quite simple. On the other hand, if the evaluation seeks the causes for success and failure, the plan must be more comprehensive. For human resources development, evaluation must have a significant role in the overall process of the human resource functions. Therefore, human resource developers need to have an understanding of the evaluation process in order to provide better quality programs (Brandenburg, 1982).

The evaluation process is concerned with how the human resource developers are meeting the objectives of the HRD program. It provides feedback to determine whether the HRD programs are relevant to the identified needs of the organization, and to the people who are concerned with meeting those needs (Hamblin, 1974; Nadler, 1980). One way for the HRD evaluator to insure that the evaluation will serve some practical purpose is to identify the clients and audiences and determine what they need to know and why they need to know it. The case for evaluation is based on the promise that people will, if they can, make changes based upon information. Therefore, decisions about programs or services will be more rational if good information is available when needed (House, 1977). Human resource developers need to know if an instructor is effective, materials are readable, facilities are adequate, and trainees are using what they learned on the job. Therefore, audiences have specific information needs when it comes to HRD and
evaluation (Bakken & Bernstein, 1982).

Good evaluation depends on the identification of important audiences and the development of a format that will both appeal to and convince the evaluation's audience (Worthen & Sanders, 1987). An evaluation must meet the information needs of its several audiences. Instructors, for example, may need information to help them improve their training skills; trainees need feedback on how well they have mastered the HRD objectives; HRD designers may need information to revise a program; and managers may need information on how HRD works on the job (Brinkerhoff, 1987).
CHAPTER III

METHODOLOGY

The purpose of this study was to investigate and explore how the focus of HRD evaluations might be influenced by different audiences and clients, and to develop recommendations to individuals and organizations involved in designing, developing, and implementing HRD evaluation based on the findings.

This chapter presents the procedures used in conducting the study. It includes the identification and description of variables, the instrument development process, the validity study and pilot test, the simulation development, hypotheses, the experimental design and data collection, and data analysis procedure. The results of this study will assist HRD evaluators to decide what types of information may be needed to satisfy the needs of various client groups.

Identification and Description of Variables

The independent variable in this study was the role of clients and audiences in organizations. Clients and audiences are individuals and groups who have interest in the HRD evaluation and receive its results. While audience may
be operationalized in a number of different ways, the context of this study (HRD) limited the operationalization of the variable. In evaluating HRD programs, the audience and clients are most likely to be the CEOs of the companies, HRD managers, those who deliver the program, and the trainees who attend the training sessions.

The dependent variables were the ratings of the importance of the phase evaluation questions specified by the CIPP model. The CIPP evaluation model provided the levels of criteria used in this study rather than the evaluation components in the Kirkpatrick or Brinkerhoff models because of the comprehensiveness of the CIPP evaluation model which covers all major aspects of evaluation. There were four types of evaluation categories involved (Context, Input, Process, and Product). Ratings of the importance of evaluation questions in each of these areas provided a comprehensive picture of the information needs for each audience.

Instrument Development

There is very little evidence to suggest which kind of evaluation questions should receive the highest priority in evaluating HRD programs either within a single organization across client types or across several organizations with the client type held constant. No investigations of this question have been performed to date, and no instrumentation exists to investigate this question. The human re-
sources evaluation questionnaire (HREQ) was developed to look at the importance of evaluation questions from the perspectives of different audiences and clients. The purpose of the HREQ was to measure the importance of a set of evaluation questions which comprehensively describe evaluation focused at each of the four phases of the CIPP model. The following section documents the procedures that were used to develop the HREQ instrument, and the scoring system developed to yield the measure of the four dependent variables.

**Instrument Plan**

The initial draft of the Human Resources Evaluation Questionnaire (HREQ) was developed from a pool of items describing information on the CIPP evaluation model. It contained forty-eight items drawn from item pools in the literature to address the four categories of the CIPP evaluation model (Martin, 1980; Nevo, 1974). The item pool developed by Nevo had been previously validated for school personnel as representative of the CIPP model. However, the Martin pool was simply a list of optional evaluation questions. There were twelve items assigned to each category of evaluation by the researcher. The rationale for selecting twelve items for each category was to give an equal weight to each evaluation category so that the in-
strument did not bias any particular phase of program development. The rating scale designed for the instrument required the participants to judge the importance of the evaluation questions on a five-point Likert-type scale. The classification of items in the initial draft is presented in Table 1 (See Appendix A for the text of the items). The instrument went through two phases of revision prior to use in this study. Phase one was a validity study, where a panel of experts determined the relation of the HREQ items to the CIPP model. Phase two was the pilot testing to give an estimate of reliability of each of the four dependent variables.

Table 1
Classification of Items in the Initial Draft

<table>
<thead>
<tr>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2,4,8,15,31,40,41,42,43,44,47</td>
<td>5,6,7,9,10,13,14,17,19,22,30,33</td>
<td>3,11,12,16,21,23,27,28,29,32,38,46</td>
<td>18,20,24,25,34,35,36,37,39,45,48</td>
</tr>
</tbody>
</table>

Panel of Experts: A Validity Study

Content validity refers to the extent to which the instrument represents the content of the program (Phillips, 1983; Brown, 1983). For the HREQ instrument, content va-
Validity refers to the extent that items can be placed with confidence within each category that relate to the CIPP evaluation model. Low content validity would mean that the HREQ items did not represent a true sample of the CIPP model or that some questions could not be put in a single category. High validity means that the HREQ items represent a good balance of all the information presented in the CIPP model categories.

A content validity panel was asked to judge how the HREQ items relate to the CIPP model categories. A letter cosigned by the researcher and the advisor was sent to each potential member of the panel of experts. The cover letter explained the purpose of the content validity and asked for their assistance. Five of the invited persons agreed to serve. One person declined because of a perceived lack of experience with the CIPP model.

The judges were selected from the Evaluation Center at Western Michigan University based on their practice with and expertise in the CIPP evaluation model. With regard to their academic credentials, the panel of experts was composed of two evaluation specialists who hold doctoral degrees with an average of ten years of experience in evaluation and three doctoral students specializing in evaluation.

The panel's task was to read each item and: (1) Check whether it fell under Content, Input, Process or Product,
or whether it fell under more than one of the categories, and, (2) Indicate whether the item was relevant from an evaluation perspective. The first task speaks to the validity of the item in each of the four dependent variables, while the second validates the items as good evaluation questions. It is possible for an item to be easily categorized and yet not be a good evaluation question (See Appendix A for a replica of the panel instrument).

An item that had less than three panel members in agreement, or that was placed in two categories, or that yielded irrelevant data in the opinion of three panel members was deleted from the instrument. Eight items were found flawed by the panel and consequently deleted from the pilot text revision of the instrument. Some items (e.g., items 31 and 41 in the Context category, items 17 and 33 in the Input category, and item 26 in the Product category) were placed in more than one category by three or more of the panel members. Other items (e.g., items 3 and 32 in the Process category, and item 20 from Product category) were found to be either unclear or irrelevant by at least three panel members.

Pilot Test: An Estimate of Reliability

A pilot test was conducted prior to administering the questionnaire in the research study. Approval of the Human
Subjects Institutional Review Board was secured prior to implementation of the pilot test. The objectives for the pilot testing were: (1) Clarification of items and simulation packets, and (2) Calculation of an estimate of reliability of each of the four dependent variables.

A group of seventeen students from a Supervision class at Western Michigan University was chosen for the pilot test of the instrumentation. They were graduate level students in Educational Leadership with an emphasis in administration not HRD. The researcher began by reading a script (See Appendix B) that asked the students' permission to participate, and provided them with a set of questions related to the content and clarification of items in the questionnaire and the simulation packets. Preliminary remarks were similar but not exactly the same as the experimentation text because the packets called for more than simple judgment of importance on the HREQ. The students were asked to make judgments of clarity of the items in addition to their judgment for item improvement. The students were asked to play an HRD role (trainee, HRD staff, HRD manager, or CEO). The directions on the questionnaire asked the students to write next to each statement if they could not understand the language or if it seemed unclear to them, and to indicate how it could be improved (See Appendix B).

The review of the frequency of the scale point used by
the pilot group found that for twenty-three items only four points of the scale (2-5) were used, and for ten items only three points of the scale (3-5) were used. The use of only a portion of the scale on thirty-three out of forty items (82% of the instrument) indicated that the participants were not differentiating among the items. Therefore, a change had to made on the scale of the questionnaire used in the experiment. The initial scale used in the pilot test was anchored by the words low importance and high importance. The scale was revised so that, 1= important, and 5= critical (See Appendix C for the final form of the instrument).

The scoring procedures for the HREQ was based on the sum of weight assigned for each response for relevant items (Brown, 1983). The overall score in each category of information was the sum of the item responses divided by ten (the number of items for each category). This procedure would result in a mean score for each category for each of the assigned roles on a five-point scale of importance.

Analysis of the data from the pilot test included both reliability estimates of the measure of the dependent variable and the discrimination power of each item. The estimate of reliability was based on the use of coefficient alpha on four of the dependent variables. Reliability refers to the consistency within the item pool (Anastasi, 1982; Cronbach, 1970). Table 2 reports the results of such anal-
ysis. The data indicated that the reliability of the instrument was very respectable for the four dependent variables.

Table 2
Reliability Coefficient for the Pilot Test

<table>
<thead>
<tr>
<th>Phase of Evaluation</th>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient Alpha</td>
<td>.81</td>
<td>.76</td>
<td>.78</td>
<td>.70</td>
</tr>
</tbody>
</table>

A number of items that proved to be confusing or did not meet the statistical criteria were revised. The refinement of the questionnaire was based on the statistical item analysis and the suggestions of the participants. An item that had lower than -.1 item total correlation in the item analysis (the discrimination index) was revised in the final form of the instrumentation. Three items that did not meet the criteria (e.g. item 33 in the Context category, items 18 and 26 in the Product category) were revised in the final form of the instrumentation. Several items were also revised based on the suggestions of the respondents.

Thus, the final form of the questionnaire was based on the information obtained from the panel of experts and the
item review. A selection was made choosing forty items, ten items for each category. The classification of items found valid by the panel of experts, and the pilot arrangement of items with the consequent new numbering system is presented in Table 3.

Table 3
Change in Instrument After Item Review

<table>
<thead>
<tr>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Product</th>
<th>Found in</th>
</tr>
</thead>
<tbody>
<tr>
<td>PANEL</td>
<td>1,2,4,8, 15,40,42, 43,44,47</td>
<td>5,6,7,9, 10,13,14, 19,22,30</td>
<td>11,12,16, 21,23,27, 28,29,38, 46</td>
<td>18,24,25, 34,35,36, 37,39,45, 48</td>
</tr>
<tr>
<td>PILOT</td>
<td>1,5,9,13, 17,21,25, 29,33,37</td>
<td>3,7,11,15, 19,23,27, 31,35,39</td>
<td>4,8,12,16, 20,24,28, 32,36,40</td>
<td>2,6,10,14, 18,22,26, 30,34,38</td>
</tr>
</tbody>
</table>

Simulation Development

Job simulation is a method that involves the construction of a task that simulates the activity in actual job situation. There are a variety of simulation techniques such as task simulation, business games, in-basket exercises, case studies, and role playing. Each exercise is designed to produce a work situation related to job dimen-
This is a simulation study, in which the participants were provided with information that simulated each audience as CEO, HRD manager, trainer, or trainee.

The simulation packets (the independent variable manipulation) went through development and pilot testing much like the measure of the dependent variables. They were developed from a combination of real job advertisements in the *Training and Development Journal* and Phillips (1983). The participants in the pilot testing confirmed the language and the clarification of the role playing simulation by answering the following questions: (1) Was the role playing simulation clear for your HRD position; (2) Was the language of the material easy to understand, and how it can be improved.

The rule for the role playing simulation revision was to seek majority agreement of the respondents answering the clarification questions. There was a majority agreement on the CEO role which indicated that the role was clear for their position. There was also a majority agreement for the HRD manager and staff roles. As far as the trainees' role, there was less agreement from the participants. They suggested the addition of more information to their role. Consequently, training history was added explaining to the trainees that they had attended two different training sessions and that they were being asked to provides us with
information that reflected the trainees' point of view. Additional information was added to each of the roles as a result of students' suggestions to include more information and directions on how to carry out their role playing simulation. The overall picture of the pilot test showed that seventy-five percent of the participants indicated that the role playing simulation was clear and understandable for their position. However, to insure that the participants in the experiment remembered their roles, two reminders of their role assignment were placed in the HREQ.

The final forms of the role playing simulation were one page long, which included their organizational role, responsibilities and their rating task. There were some common elements in each of the simulations such as the organization and the subject. The differences were in the organizational role and responsibilities (See Appendix D for copies of simulation material).

Design and Data Collection

This simulation study was conducted in graduate level classrooms which focus on HRD. The reason for not selecting real role incumbents was no previous research in the area. The researcher believed that through simulation we can get as close to actual HRD experience at the decision-making level as possible without leaving the classroom, by developing a scenario in which all organizations are iden-
tically equal, thereby eliminating potential across organizational confounds. Common organizational setting is critical in this sort of study. If organizational size, profits, and other financial comparison were used for evaluations in real organizations, then no group would have had a common basis on which to judge evaluation questions.

Operational Procedures

The target population were clients and audience who are responsible for judgments to provide HRD programs with information for developing their human resources as part of the developmental process. The participant pool for the study were five HRD classes offered through the Educational Leadership Department at Western Michigan University. The courses selected for the study were all the HRD courses offered in Winter 1993 schedule of classes; two sections of Introduction to HRD, a section of Needs Assessment, and two sections of Project Management. The reason for selecting only HRD courses was due to the fact that students who usually enroll either hold or seek an HRD position. Thus, the participant pool had some grounding in the requirements of each role.

A letter was sent to four professors asking them for permission to include their students in the study (See appendix E). After a three to five day period, a telephone
call was made to each professor to confirm agreement and to arrange a time and place for conducting the study. All professors agreed to participate. The criterion of HRD classes limited the potential pool of participants, but was necessary to protect the validity of the study. Four packet types consisted of a questionnaire and a role statement. Each packet was printed in a different color to allow the coding of the independent variable.

The packets of the questionnaire and role statement were arranged in random order by role prior to initiating the study in each class. The researcher used a deck of cards and assigned each role a suit, shuffled the cards and began randomly ordering the simulation roles. For instance, hearts were assigned the CEO category, each time a heart was drawn from the deck the CEO packet was put into the stack. Before distributing the questionnaire from the stack to the students, a script was read asking their permission to participate in the study and to confirm the anonymity of the data. The time task for reading the simulation role and filling out the questionnaire was approximately twenty minutes.

Hypotheses

In this study the following hypotheses were tested:

Hypothesis #1: The role of the audience in organiza-
tions influences the perceived importance of evaluation questions differently if the information is categorized as Context information in the CIPP evaluation model.

Hypothesis #2: The role of the audience in organizations influences the perceived importance of evaluation questions differently if the information is categorized as Input information in the CIPP evaluation model.

Hypothesis #3: The role of the audience in organizations influences the perceived importance of evaluation questions differently if the information is categorized as Process information in the CIPP evaluation model.

Hypothesis #4: The role of the audience in organizations influences the perceived importance of evaluation questions differently if the information is categorized as Product information in the CIPP evaluation model.

Data Analysis Procedure

To determine the importance of the evaluation questions related to HRD programs, the participants were asked to determine the importance of each question on a five-point scale where, 1 = importance, and 5 = critical, without being cued as to the relevant phase of evaluation work. The scoring procedures for the HREQ was depended on the category of the CIPP model for each item and the weight assigned for each response (Brown, 1983). The overall score
sum of item responses in each category of information was divided by ten (the number of items for each category) that would result in a mean score for each variable on a five-point scale of importance. Differences could then be observed regarding the importance of the various categories of information.

One-way analysis of variance was the statistical procedure that was used to test the null form of four major hypotheses. The purpose of using this statistical procedure was to analyze the four levels of the independent variable separately and collectively in relation to the dependent variable; the importance of phase evaluation questions (Context, Input, Process, and Product).

The test statistic that was used in ANOVA is the F ratio of between groups variation and within group variation using a .05 level of significance. Rejecting the null hypothesis indicates that there is a significant difference among the sample means but does not identify which of the four means are different (Hinkle, Wiersma, & Jurs, 1988). LSD multiple comparison procedure was applied to determine which means were different from each other.
CHAPTER IV

RESULTS OF THE STUDY

This chapter presents the results of the study, including an overview of the study, the instrument characteristics, testing of the hypotheses, and a correlational analysis. With regard to the central problem of the study, no prior investigation has examined the influence that organizational roles in the HRD field have on the perceived importance of evaluation questions in the four categories of the CIPP model or in any other models. Most HRD evaluations focus on the reaction sheet of the trainees suggesting a common value for program success. Whereas evaluation theory focuses more on how audiences perceive the information needed. The objective of this study was to investigate whether the type of evaluation questions regarding the HRD effort has different importance to audiences depending upon their role in the organization.

The findings of the study are presented based on the testing of the four hypotheses. Each hypothesis was tested with data obtained from four different audiences (CEO, HRD manager, HRD staff, and trainee). The independent variable in the study was the role of audiences and clients in orga-
nizations. The dependent variables were the ratings of importance of the phase evaluation questions.

The population of the study consisted of the audience and clients who are responsible for evaluating HRD programs. The data were collected from a pool of five University classes. The sample consisted of seventy-three graduate students enrolled during the 1993 Winter Semester at Western Michigan University in the five HRD classes. These students either hold or seek an HRD position. There were twenty-five students in the first Project Management class, five in the second one, seventeen students in the first Introduction to HRD class, eighteen in the second one, and eight students in Needs Assessment class. The small number in two of the classes influence the overall number of participants, thus restricted the sensitivity of the design to differentiate between the role groups. The design was based on an estimate of twenty students per class or one hundred twenty participants. The resultant sample size of 60% of the potential pool made the design less sensitive to differences between the groups.

Instrument Characteristics

The instrument used in the research resulted from extensive development and pilot testing. The instrument consisted of forty items, ten items for each category. The panel of experts confirmed the validity of items related to
the CIPP evaluation model. The pilot test confirmed that
the items were clearly stated and gave an estimate of reli-
ability.

The human resources evaluation questionnaire (HREQ)
yielded four scores for each participant. The scores were
determined by adding the items associated with each depen-
dent variable and dividing by ten. Table 4 shows the char-
acteristics of each variable across all participants.

Table 4

<table>
<thead>
<tr>
<th></th>
<th>Overall Mean</th>
<th>Low Score</th>
<th>High Score</th>
<th>SD</th>
<th>Coeff. Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>3.51</td>
<td>1.60</td>
<td>4.90</td>
<td>.62</td>
<td>.70</td>
</tr>
<tr>
<td>Input</td>
<td>3.45</td>
<td>1.80</td>
<td>4.60</td>
<td>.61</td>
<td>.74</td>
</tr>
<tr>
<td>Process</td>
<td>3.24</td>
<td>1.60</td>
<td>4.60</td>
<td>.65</td>
<td>.77</td>
</tr>
<tr>
<td>Product</td>
<td>3.67</td>
<td>1.70</td>
<td>4.80</td>
<td>.59</td>
<td>.70</td>
</tr>
</tbody>
</table>

Table 4 presents the general view of the way partici-
pants perceived the evaluation information. The scale of
measurement was on five-point Likert scale where
1 = important, and 5 = critical. The mean score for the
Product category was higher than the mean score of the
other categories with ratings ranging from 1.7 to 4.8 and a
standard deviation of .59. For normal distributions sixty-
eight percent of the ratings from participants fall within one standard deviation from the mean. Thus, a majority of responses for the participants in this study were between 3.08 and 4.26, all clearly above the midpoint of the scale. On the other hand, the mean for the Process category was lower than the means of the other categories, with ratings ranging from 1.6 to 4.6 and a standard deviation of .65. That was an indication that the Process evaluation information was perceived to be relatively less important by participants. Recall that reaction forms or process information is the most common form of HRD evaluation.

As far as the reliability of the instrument, the Coefficient Alpha was the statistical procedure used to determine the reliability of the instrument in this study. The instrument was found to be very reliable for the four evaluation categories with the highest of .77 for the Process category.

Test of Hypotheses

The data were analyzed primarily through the use of the SPSS statistical program. The data were examined by using the reliability analysis, correlation coefficient, and analysis of variance. The following presents the statistical analyses for the four hypotheses.

Hypothesis #1: The role of the audience in organizations influences the perceived importance of evaluation
questions differently if the information is categorized as Context information in the CIPP evaluation model.

The Context category provides evaluation information related to goals, objectives, and needs. The null hypothesis was tested by ANOVA one-way analysis of variance to see if there was a difference between the four organizational roles regarding the perceived importance of Context information. The ANOVA yielded an F ratio of .486 with 3 and 69 degrees of freedom and a probability of .69. Table 5 presents the results of the analysis of variance.

Table 5
Analysis of Variance Summary for the Context Category

<table>
<thead>
<tr>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>3</td>
<td>.571</td>
<td>.190</td>
<td>.486</td>
</tr>
<tr>
<td>Within</td>
<td>69</td>
<td>27.00</td>
<td>.391</td>
<td></td>
</tr>
</tbody>
</table>

Based on these data, it was not possible to reject the null hypothesis. Thus, the data from the four organizational roles did not support the first hypothesis. In other words, no evidence was found for differences between the perceived importance of Context evaluation questions based on the organizational roles of the audience.

Hypothesis #2: The role of the audience in organizations influences the perceived importance of evaluation
questions differently if the information is categorized as Input information in the CIPP evaluation model.

The Input category provides evaluation information related to design, plans, and strategies. To test the second null hypothesis, the same procedures were followed as for the first hypothesis. Table 6 presents the summary of the analysis of variance.

Table 6
Analysis of Variance Summary for the Input Category

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>3</td>
<td>.630</td>
<td>.210</td>
<td>.553</td>
<td>.65</td>
</tr>
<tr>
<td>Within</td>
<td>69</td>
<td>26.21</td>
<td>.379</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 6, no significant relationship was found. The results yielded an F ratio of .553 with 3 and 69 degree of freedom and probability of .65. In other words, there was no significant difference between the perceived importance of information in the Input evaluation category based on the organizational roles.

Hypothesis #3: The role of the audience in organizations influences the perceived importance of evaluation questions differently if the information is categorized as Process information in the CIPP evaluation model.
The Process category provides evaluation information related to implementation, materials, and program activities. To test the third null hypothesis, the same procedures were followed as for the other hypotheses. Table 7 presents the summary of the analysis of variance regarding the Process category.

Table 7
Analysis of Variance Summary for the Process Category

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>3</td>
<td>.343</td>
<td>.114</td>
<td>.260</td>
<td>.85</td>
</tr>
<tr>
<td>Within</td>
<td>69</td>
<td>30.38</td>
<td>.440</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 7, no significant relationship was found. The results yielded an F ratio of .260 with 3 and 69 degree of freedom and probability of .85. In other words, there was no significant difference between the perceived importance of information in the Process evaluation category based on the organizational roles.

Hypothesis #4: The role of the audience in organizations influences the perceived importance of evaluation questions differently if the information is categorized as Product information in the CIPP evaluation model.

The Product category provides evaluation information related to outcomes or results, achievement of objectives,
and impact. To test the fourth null hypothesis, the same procedure was followed as for the other hypotheses. Table 8 presents the summary of the analysis of variance regarding the Product category.

Table 8
Analysis of Variance Summary for the Product Category

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>3</td>
<td>.173</td>
<td>.057</td>
<td>.159</td>
<td>.92</td>
</tr>
<tr>
<td>Within</td>
<td>69</td>
<td>24.98</td>
<td>.362</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 8, no significant relationship was found. The results yielded an F ratio of .159 with 3 and 69 degree of freedom and probability of .92. In other words, there was no significant difference between the perceived importance of information in the Product evaluation category based on the organizational roles.

Correlational Analysis

As discussed in the instrument characteristics, most rating were positive, i.e., participants valued all evaluation information, but the question was do they differentiate between the types of information. While there was no significant difference among the four organizational roles to support the research hypotheses, the correlation coeffi-
cients of the four categories yielded interesting results. There was a high correlation among all ratings of evaluation questions categories which suggests that the study participants perceived all information as related. Table 9 presents the correlation coefficients regarding the four type of information. Given that there was variance in the assessment of importance, these correlations are particularly interesting. Persons who rated the need of information low in one category were likely to exhibit low information needs in all categories.

Table 9
Correlation Coefficients for the Four Types of Information

<table>
<thead>
<tr>
<th></th>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>.73*</td>
<td>.69*</td>
<td>.71*</td>
<td></td>
</tr>
<tr>
<td>Input</td>
<td></td>
<td>.68*</td>
<td>.59*</td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td></td>
<td></td>
<td>.64*</td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .0001

The correlation between the Context and Input category questions was the highest with a correlation of .73, while the correlation between the Input and Product questions had the lowest correlation among all categories .59. All of
the correlations were statistically significant with \( p < .0001 \).

Summary of the Chapter

The results of the data analysis did not support the hypotheses. In other words, no significant differences were found with respect to the importance of evaluation questions with the four categories of the CIPP model based on the organizational role. There were significant high correlations between the four evaluation categories indicating that participants perceived all information as equally important irrespective of the CIPP model categories. Persons who found any information important were likely to find all information important.

Participants value of information is not determined by the type of information. Ratings of the importance of any kind of evaluation questions were co-variate with all other kinds of evaluation questions. Each of the dependent measures had means above the midpoint of the five-point scale indicating a positive view of evaluation questions. Finally, the data analysis showed a respectable reliability for the four categories in the HREQ instrument.
CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

This study was designed to investigate the importance of the focus of human resources evaluation questions related to different audiences and clients. None of the four research hypotheses was supported. Evidence was found of a strong interrelationship among the four dependent variables, suggesting organizational participants rated questions across the four categories in a systematic fashion. Additionally, the mean rating for the process data (the data commonly obtained in current HRD evaluations) were the lowest value.

This chapter presents the conclusions of the study. First, the current issues and problems that were discussed in early chapters are reviewed. Second, the limitations which includes the strengths and weaknesses of the study are presented. Third, recommendations for future studies are suggested. Finally, the conclusions are presented.

Discussion of Findings

The problems of HRD evaluation identified in the early chapters provided the rationale for conducting this research. HRD evaluators tend to look almost exclusively at
participants' reactions and learning, suggesting a common value for their efforts. They have utilized the participant evaluation in an attempt to gauge the effectiveness of their programs. However, the evaluation literature focuses on audience and client needs in program evaluation. Both of these contentions were brought into question by the results of this study.

The goal of HRD evaluators should be to develop an evaluation strategy which reflects the expectations of various audiences rather than simply process evaluation. The evaluation strategy is a way of determining the worth and merit of the HRD intervention to the different audiences involved. As was previously mentioned, most HRD evaluations do not follow logical steps with respect to the design and strategies, so that they can fulfill the information needs of different audiences.

The CIPP evaluation model presented in this study provided the strategy for carrying out this HRD evaluation research. It is a decision-oriented model that focuses on providing information to decision makers. The four types of evaluation are based on four types of decisions in education: objectives, design, delivery, and impact. The model provides a balanced view of evaluation, and has proven to be useful and practical for educational evaluation. The model implies differences among the evaluation...
questions. Yet the current data show no differences when the questions are posed generally.

The results of the current study conducted within an HRD framework, are interestingly similar to those reported by Nevo (1974) in an educational environment. Nevo (1974) did find that the different roles in education have either influenced the importance or the availability of evaluation information. But his results were not across all roles.

The results presented in the previous chapter, even with no support for the research hypotheses, indicates that HRD evaluators need to rethink their current emphasis on process evaluation, expanding their evaluation to include context, input, and product information.

In discussing the implications of the results, it is important to note the unique aspects of the population from which the sample was drawn. All participants were in a graduate level HRD program and role played the various organizational audiences and clients. Failure to use real incumbents as participants may be the main weakness of the study.

Limitations of the Study

Certain limitations of the study should be noted as not to mislead others when considering similar research for HRD evaluation. The limitations will be discussed in terms
of the design of the study, the independent variable, and the dependent variables.

This study has merit and can serve as a useful guide for others interested in HRD evaluation due to the fact that it was based upon information collected from several simulated HRD audiences, i.e., CEOs, managers, trainers, and trainees. The generalization of the results, however, are limited by a number of issues. Limited generalization is common practice in research and should not be seen as a fatal flaw.

First, this study was limited by the small sample size. The small size of the sample produced a very high Type II error probability of .72 (Cohen, 1969). While the Type I error rate is determined exclusively by the researcher, the Type II rate depends upon the size of the sample among other factors. The sensitivity desire and the variance of the instrument in addition to other factors, would have produced respectable power with a larger sample size.

Although restriction of the sample pool from HRD courses protected the validity of the study, due to the familiarity of the subjects with the organizational roles they were asked to assume, this procedure yielded too few subjects.

Second, the simulation materials associated with organizational roles might have been inadequate. That is, the
participants may have not reacted to the evaluation ques-
tions as those who actually hold those positions in organi-
zations. When participants were told that they were in a
position to evaluate programs they might have perceived all
evaluation information as important because of the simu-
lated nature of the study. In contrast to real organiza-
tional members who must actually respond to use the evalua-
tion data for specific purposes. Therefore, it is unknown
whether real organizational audiences and clients would
have provided ratings that would have resulted in the high
inter-correlations between the CIPP model measures as ob-
served with the current participants. However, a design
with real role incumbents would have to control for contam-
inating variables of company size, HRD history, type of or-
ganization, etc., in order to produce valid results, a dif-
culty addressed earlier. The trade-off between sources
of contamination which limited generalization must be made
in the context of each piece of research.

Third, perhaps the dependent variables were not con-
sistent with the way real people think about evaluation.
This study was based upon the application of the CIPP eval-
uation model. The CIPP evaluation model was designed for
decision making in educational setting. The decision situ-
atations in the role simulations were not anchored to a spe-
cific program or to a specific phase of the program.
Therefore, the HRD people may perceived all categories as
important for their evaluation of various phases of a pro-
gram.

Finally, the changing of the scale measure of depend-
ent variables in the instrument might be one of the weak-
nesses of this study. In the pilot version of the instru-
ment, the scale ranged from low importance to high impor-
tance. In the participant survey the scale ranged from im-
portant to critical. The use of full scale increased the
variance and thus, may have masked differences. The fre-
quency of the scale point used in the participant survey
found that thirty-nine items used full scale and one item
used four points. Thus, the change in scale clearly intro-
duced more variability in participant ratings. But, the
instrument had very good reliability that could serve as
guide to develop evaluation plans.

Recommendations for Future Research

The major finding of the study was the discovery of no
evidence that HRD organizational audiences and clients per-
ceived the importance of the evaluation information differ-
ently as it is categorized in the CIPP model. However, re-
lationships of the data were discovered. The data showed
that there were high correlations between the four cate-
gories in the CIPP model, indicating that all evaluation
categories are perceived as systematically similar. The
findings suggests that the HREQ instrument has high valid-
ity and reliability and thus could be utilized as planning tool for HRD evaluation.

Future research is needed to determine what type of information HRD evaluators should include and present to organizational audiences and clients. This is particularly important because HRD evaluators currently focus on process evaluation, yet this was perceived by the current participants as the least important of the four CIPP categories. Therefore, it would be useful to replicate and refine this study.

Future studies should include data collection from a larger sample. Studies with a larger samples may provide additional information on HRD evaluation.

Future studies should also consider the substitution of the CIPP evaluation model. The CIPP model has not been operationalized or tested for HRD evaluation. The CIPP model was chosen for this study for its ability to categorized information and the comprehensiveness of that information. While the outcomes for the CIPP categories did not provide significant data for the study, the information appeared to be moderately important for all audiences. The use of another valid model may provide significant differences based on HRD audiences.

Future studies should include data collected from HRD incumbents employed by organizations with similar characteristics. While the results of this study were based upon
information provided in the role playing simulation, further research is needed with HRD incumbents. Such a design would increase the generalizability of outcomes to the population.

Conclusion

The purpose of the study was to investigate and explore the extent to which the importance of evaluation questions differ across HRD audiences and clients. As with most research, this study identified additional questions which should be investigated further.

This study provided an orientation to the evaluation literature especially with respect to concepts of the formative/summative evaluation, the use of the terms merit/worth, and impact evaluation. It also introduced the HRD evaluation models that have some deficiencies with regard to evaluation terminology and practice. Therefore, this study was important in its attempts to fill some of the gaps in HRD evaluation and the types of questions HRD evaluators work with by combining the evaluation literature and the HRD practice.

This study was an attempt to find whether evaluations have a different purpose for different audiences within the HRD context. It used a unique method of data collection through the use of role playing simulation. Investigations using HRD incumbents should be conducted to examine the or-
ganizational roles and the perceived importance of evaluation information.

The results yielded no support for a relationship between the organizational roles and the perceived importance of different types of information. However, the results suggest that HRD role participants perceived all information as important when its categorized in the CIPP model. Further investigation with different model is recommended.

There were some limitations reported in this study. In spite of its limitations such as sample size, and the role simulation, the study provided new knowledge in HRD evaluation. The questionnaire developed for this study might be a significant tool for planning HRD evaluations because it clearly puts forth a series of questions which might be used in evaluation.
Appendix A

Letter of Request to the Content Validity Panel
Because of your expert knowledge and understanding of the CIPP Evaluation Model, I would like you to serve as a member of the Content Validity Panel to review the instrumentation for my doctoral dissertation research in the Educational Leadership Department. By assisting me, you can contribute to our knowledge regarding the evaluation of Human Resource Development programs.

Your assistance in validating the instrument will be greatly appreciated. The task should take no longer than twenty minutes. The results of the study will be available to you, if you are interested. Please indicate your desire to review the results in your response.

I will call you in a few days. If you agree to participate I will set up an appointment with you and deliver the questionnaire to you at that time. Thank you

Sincerely,

Fahad Al-Ruwaished
Doctoral Candidate

Mary Anne Bunda, Ph.D.
Dissertation Chair
Human Resources & Evaluation Questionnaire

This questionnaire is designed to determine your opinion about the types of questions that are relevant to the CIPP evaluation model. Following is a list of statements describing various types of information that might be useful for HRD practitioners to evaluate their efforts. Please read each statement and express your opinion regarding the CIPP evaluation model. For each statement of information, circle the letter facing the question in a category (Context, Input, Process, or Product) only if it is a good evaluation question. Any item that would be placed in two categories or that in your opinion would yield irrelevant data should be crossed out.

Aspect of Evaluation

1. Information about trainees needs, that would help to determine how well the program is responding to these needs. C I P P

2. Information about any prerequisites required of trainees or staff? C I P P

3. Information on how the prerequisites effect the selection of trainees. C I P P

4. Information about professional aspirations of trainees that would help to prepare them better for their future roles. C I P P

5. Information about what was included in the planned program? C I P P
Aspect of Evaluation

6. Information on how did this information compare to similar programs in other institutions? C I P P

7. Information about what materials were specified during planning and development? Were these materials a major concern of the program? C I P P

8. Information about what group was the program targeted? How were these groups identified? C I P P

9. Information on how many participants was the program designed to serve? C I P P

10. Information about what personnel resources were available for the program? How were the personnel chosen? C I P P

11. Information about how well were the physical facilities suited to the needs of the program? C I P P

12. Information about the appropriateness of learning resources provided to trainees during the program. C I P P

13. Information about the merit of various evaluation methods, that would improve the evaluation system of the HRD program. C I P P

14. Information about the objectives of the program? How these objectives stated/ Do they cover all aspects of the program? Can they be applied? C I P P

15. Information about staff needs and priorities, that would help acquire funds for staff development. C I P P
Aspect of Evaluation

16. Information about what activities occur during instruction? C I P P

17. Information on how are the activities tied to the program? C I P P

18. Information about trainees achievement, that would help to evaluate the effectiveness of the HRD staff. C I P P

19. Information about how the materials related to the instructional objectives of the program? C I P P

20. Information about the objectives of various subjects that been offered, that would help assess the merit of the HRD goals. C I P P

21. Information about how does the teaching style of each trainer fit into the program? What seem most effective for the program? C I P P

22. Information about the relative importance of program objectives, that would help to determine what curriculum is best suited to meet these objectives. C I P P

23. Information about what outside classroom activities are included as a part of the program/ How are these outside classroom activities related to in classroom activities? C I P P

24. Information about how are trainees evaluation results reported? Are these reports verbal or written? C I P P

25. What information do clients receive on trainee evaluation? C I P P
Aspect of Evaluation

26. Information on how is evaluation communicated?  C I P P

27. Is there any form of instructor evaluation in the program? How does it occur? How are the results used?  C I P P

28. Information on what type of supervision is provided by superiors?  C I P P

29. Information on how do instructors react to supervision? Is the data used as a means of instructor improvement or as an assessment of merit?  C I P P

30. Information about priorities in staff development, that would help to develop a better in service training program.  C I P P

31. Do assessments of the program and its performance occur periodically?  C I P P

32. Information on what level do assessment occur? Is it voluntary or required?  C I P P

33. Information about the strengths and weaknesses of the staff to help assess their competence as educators.  C I P P

34. Information about the achievements of each trainee during the process of his/her learning, that will help to direct them to the appropriate instructional materials that will remedy the gap.  C I P P
Aspect of Evaluation

35. Information about trainees' achievement, that would allow trainer to provide them with meaningful feedback regarding their progress in studies.

36. Information about what gains in knowledge have occurred? How rapidly did these gains occur? Which of these gains fulfilled program objectives?

37. Information about what attitudes have changed as a result of the program?

38. Information on how did attitudes effect program development and operation?

39. Information about how do program participants and outside specialists assess various parts of the program? What is their judgment on program worth?

40. Information about the main needs of trainees that would help to improve the ability to respond to the needs of the individual trainee.

41. Information about the academic aptitudes of trainees that would help to assess their educational accomplishments.

42. Information about the reasons trainees take HRD programs, that would help to better understand trainees motivation.
Aspect of Evaluation

43. Information about the trainees' main reason for taking courses that would help to assess their educational motivation.

44. Information about the relative importance of the objectives of the HRD programs, that would help to develop the final evaluation of the programs.

45. Information about the specific achievement of each trainee, that could serve as a basis for individual assignment.

46. Information about the efficiency in using AV equipment, that would help the utilization of facilities.

47. Information about HRD priorities in the community, that would help to increase the responsiveness to community needs.

48. Information about what new technical skills were acquired? To which area of the program were they related?

Thank You For Your Time!
Appendix B

Pilot Survey Instrument

74
Western Michigan University
Kalamazoo, Michigan 49008

Dear <<Title>> <<LName>>:

My name is Fahad Al-Ruwaished. I am a Doctoral Candidate at Western Michigan University in the Educational Leadership program with an emphasis in Human Resource Development (HRD). I have chosen to research the importance of the evaluation questions and the information related to different clients and audiences.

I would like to ask your permission to include students from your class EDLD <<>> for the pilot test of the instrumentation. The students in your class will be assigned an HRD role (Trainees, HRD staff, HRD managers, or CEO). They will then fill out a questionnaire about the importance of evaluation questions.

At no point will any individual be identified. Thus, the information they provide will be anonymous. I would appreciate it if you could give me twenty minutes from your class during the Winter Semester to conduct this study.

I will contact you soon to discuss this with you.

Sincerely,

Fahad Al-Ruwaished
Doctoral Candidate

Mary Anne Bunda, Ph.D.
Dissertation Chair
Dear Student:

My name is Fahad Al-Ruwaished. I am a Doctoral Candidate at Western Michigan University in the Educational Leadership program with an emphasis in Human Resources Development and my advisor is Dr. Mary Anne Bunda 7-3031. I have chosen to research the importance of the evaluation questions and the information related to different clients and audience. I would like to ask your permission to participate in the pilot test of my study.

At no point in your response will any individual be identified and you have the right not to participate. Thus, the information you provide will be anonymous.

Before distributing the questionnaire, the class will be divided into four groups. Each group will be assigned an HRD role (Trainee, HRD staff, HRD manager, or CEO). First, I would like you to read your simulation role. Second, you will fill out a questionnaire about the importance of evaluation questions according to your role. The task will take no longer than twenty minutes of your time. If you have any question, please feel free to ask.

Thank You
Clarification of Role Playing Simulation

1. Was the Role Play Simulation material clear for your HRD role?
   Yes  O  No  O

   If you answer no. Please indicate where was the problem
   and how in your opinion it can be improved:

2. Was the Language of the material easy to understand?
   Yes  O  No  O

   If you answer no. Please indicate which part of the
   simulation language was hard for you to understand:

3. What was your overall concern about the Role Playing
   Simulation?
This questionnaire is designed to depict your opinion on the importance of evaluation questions and information regarding the HRD & evaluation. Following is a list of statements describing various types of information that might be useful for HRD practitioners to evaluate their efforts. Please for each statement of information:

A) Indicate the relative importance of each statement to you as it applies to your HRD programs.
B) Write next to each statement if you can not understand the language of the question.
C) Circle words that you are not familiar with or it seems unclear to you.

<table>
<thead>
<tr>
<th>Aspect of Evaluation</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information about trainees needs, that would help to determine how well the program is responding to these needs.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information about trainees achievement, that would help to evaluate the effectiveness of the HRD staff.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information about what was included in the planned program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information about how well were the physical facilities suited to the needs of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information about any prerequisites required of trainees or staff?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information about how are trainees evaluation results reported? Are these reports verbal or written?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Information on how did this program compared to similar programs in other institutions?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Information about the appropriateness of learning resources provided to trainees during the program.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Information about professional aspirations of trainees, that would help to prepare them better for their future roles.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. What information do clients receive on trainee evaluation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Information about what materials were specified during planning and development? Were these materials a major concern of the program?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Information about what activities occur during instruction?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Information about what group was the program targeted? How were these groups identified?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Information about the achievements of each trainee during the process of his/her learning, that will help to direct them to the appropriate instructional materials that will remedy the gap.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Information on how many participants was the program designed to serve?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Information about how does the teaching style of each trainer fit into the program? What seem most effective for the program?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Information about staff needs and priorities, that would help acquire funds for staff development.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low Importance</th>
<th>High Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Low</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>18. Information about trainees' achievement, that would allow trainer to provide them with meaningful feedback regarding their progress in studies.</td>
<td>1</td>
</tr>
<tr>
<td>19. Information about what personnel resources were available for the program? How were the personnel chosen?</td>
<td>1</td>
</tr>
<tr>
<td>20. Information about what outside classroom activities are included as a part of the program? How are these outside classroom activities related to in classroom activities?</td>
<td>1</td>
</tr>
<tr>
<td>21. Information about the main needs of trainees that would help to improve the ability to respond to the needs of the individual trainee.</td>
<td>1</td>
</tr>
<tr>
<td>22. Information about what gains in knowledge have occurred? How rapidly did these gains occur? Which of theses gains fulfilled program objectives?</td>
<td>1</td>
</tr>
<tr>
<td>23. Information about the merit of various evaluation methods, that would improve the evaluation system of the HRD program.</td>
<td>1</td>
</tr>
<tr>
<td>24. Is there any form of instructor evaluation in the program? How does it occur? How are the results used?</td>
<td>1</td>
</tr>
<tr>
<td>25. Information about the reasons trainees take HRD programs, that would help to better understand trainees motivation.</td>
<td>1</td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Low Importance</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>26. Information about what attitudes have changed as a result of the program?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>27. Information about the objectives of the program? How these objectives stated? Do they cover all aspects of the program? Can they be applied?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>28. Information on what type of supervision is provided by superiors?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>29. Information about the trainees' main reason for taking courses that would help to assess their educational motivation.</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>30. Information about how do program participants and outside specialists assess various parts of the program? What is their judgment on program worth?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>31. Information about how the materials related to the instructional objectives of the program?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>32. Information on how do instructors react to supervision? Is the data used as a means of instructor improvement or as an assessment of merit?</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>33. Information about the relative importance of the objectives of the HRD programs, that would help to develop the final evaluation of the programs.</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Low</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>34. Information about the specific achievement of each trainee, that could serve as a basis for individual assignment.</td>
<td>1</td>
</tr>
<tr>
<td>35. Information about the relative importance of program objectives, that would help to determine what curriculum is best suited to meet these objectives.</td>
<td>1</td>
</tr>
<tr>
<td>36. Information on how did attitudes affect program development and operation?</td>
<td>1</td>
</tr>
<tr>
<td>37. Information about HRD priorities in the community, that would help to increase the responsiveness to community needs.</td>
<td>1</td>
</tr>
<tr>
<td>38. Information about what new technical skills were acquired? To which area of the program were they related?</td>
<td>1</td>
</tr>
<tr>
<td>39. Information about priorities in staff development, that would help to develop a better in-service training program.</td>
<td>1</td>
</tr>
<tr>
<td>40. Information about the efficiency in using AV equipment, that would help the utilization of facilities.</td>
<td>1</td>
</tr>
</tbody>
</table>

Thank You For Your Time!
Appendix C
Survey Instrument
This questionnaire is designed to depict your opinion on the importance of evaluation questions and information regarding the HRD & evaluation. Following is a list of statements describing various types of information that might be useful for HRD practitioners to evaluate their efforts. For each statement of information, please indicate the relative importance for the statement as it applies to your HRD programs.

<table>
<thead>
<tr>
<th>Aspect of Evaluation</th>
<th>Important</th>
<th>Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Information about trainees' needs, that would help to determine how well the program is responding to these needs.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>2. Information about trainees' achievement, that would help to evaluate the effectiveness of the HRD staff.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>3. Information about what was included in the planned program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>4. Information about how the physical facilities were suited to the needs of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>5. Information about any prerequisites required of trainees or staff?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>6. Information about how are trainees' evaluation results were reported? Were these reports verbal or written?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>7. Information about how this program compares to similar programs in other institutions?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Important</td>
<td>Critical</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>8. Information about the appropriateness of learning resources provided to trainees during the program.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>9. Information about professional aspirations of trainees, that would help to prepare them better for their future roles.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>10. Information about how trainees evaluated the programs.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>11. Information about what materials were specified during planning and development? Were these materials a major concern of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>12. Information about what activities occur during instruction?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>13. Information about which group the program was targeted at? How were these groups identified?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>14. Information about the achievements of each trainee during the process of his/her learning, that will help to direct them to the appropriate instructional materials.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>15. Information on how many participants was the program designed to serve?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>16. Information about how the teaching style of each trainer fits into the program? What seem most effective for the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>

Remember your role is [Name of role located here]
<table>
<thead>
<tr>
<th>Aspect of Evaluation</th>
<th>Important</th>
<th>Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>17. Information about staff needs and priorities, that would help acquire funds for staff development.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>18. Information about trainees' achievement.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>19. Information about what personnel resources were available for the program? How were the personnel chosen?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>20. Information about what outside classroom activities are included as a part of the program? How are these outside classroom activities related to in classroom activities?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>21. Information about the main needs of trainees that would help to improve the ability to respond to the needs of the individual trainee.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>22. Information about what gains in knowledge have occurred? How rapidly did these gains occur? Which of these gains fulfilled program objectives?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>23. Information about the merit of various evaluation methods, that would improve the evaluation system of the HRD program.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>24. Is there any form of instructor evaluation in the program? How does it occur? How are the results used?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>25. Information about the reasons trainees take HRD programs, that would help to better understand trainees motivation.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Important</td>
<td>Critical</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>26. Information about which work attitudes have changed as a result of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>27. Information about the objectives of the program? How these objectives are stated/ Do they cover all aspects of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>28. Information on what type of supervision is provided by superiors?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>29. Information about the trainees' main reason for taking courses that would help to assess their educational motivation.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Remember your Role is [Name of role located here]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Information about how program participants and outside specialists assess various parts of the program? What is their judgment on program worth?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>31. Information about how the materials related to the instructional objectives of the program?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>32. Information on how instructors react to supervision? Is the data used as a means of instructor improvement or as an assessment of merit?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>33. Information about the relative importance of the objectives of the HRD programs.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>Aspect of Evaluation</td>
<td>Important</td>
<td>Critical</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>34. Information about the specific achievement of each trainee, that could serve as a basis for individual assignment.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>35. Information about the relative importance of program objectives, that would help to determine what curriculum is best suited to meet these objectives.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>36. Information on how did attitudes affect program development and operation?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>37. Information about HRD priorities in the community, that would help to increase the responsiveness to community needs.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>38. Information about what new technical skills were acquired? To which area of the program were they related?</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>39. Information about priorities in staff development, that would help to develop a better in-service training program.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>40. Information about the efficiency in using AV equipment, that would help the utilization of facilities.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>

Thank You For Your Time!
Appendix D

Simulation Role for Participants
Role Playing Simulation

Organizational Role: C E O
Organization: Corporation of a Fortune 500 located in North America.
Subject: HRD Evaluation

Responsibility: Take the challenge of leading a Corporation of more than five hundred employees including Human Resources Department (HRD). The HRD is responsible for all elements of the company's human resources programs with its business strategy including: recruitment, orientation, training and development at all levels.

Rating Task: In 1992 the HRD department has presented a variety of programs for top management, middle management, and new employees. The purpose of the programs was to develop managerial skills and improve efficiency.

As you know the company is cutting 30% of the budget for human resources department next year. Therefore, the new philosophy of the department is to evaluate the HRD programs at the highest level possible, without allowing the evaluation process cost to exceed potential benefits. Practical application of this philosophy ensures efficient evaluation utilization. The evaluation process is designed to: (1) Determine the impact of training programs on business objectives, and (2) Improve Corporate Training programs and their impact on employee development.

Attached is an evaluation questionnaire designed to depict your opinion related to our HRD programs and what type of information is more important from your perspective.
Role Playing Simulation

Organizational Role: HRD Manager
Organization: Corporation of a Fortune 500 located in North America.
Subject: HRD Evaluation

Responsibility: Take the challenge of leading Human Resources Department (HRD) in aligning all elements of the company's human resources programs with its business strategy including: recruitment, orientation, training and development at all levels. Develop long and short-term strategies covering these areas. You have about twenty staff in your department.

Rating Task: In 1992 the HRD department has presented over two hundred programs in such areas as retirement planning, computer capability, and stress management for top management, middle management, and new employees. The purpose of the programs was to develop managerial skills and to improve efficiency.

As you know the company is cutting 30% of the budget for the human resources department next year. Therefore, the new philosophy is to evaluate our HRD programs without allowing the evaluation process cost to exceed potential benefits. Practical application of this philosophy ensures efficient evaluation utilization. The evaluation process is designed to: (1) Determine the impact of training programs on business objectives, and (2) Improve Corporate Training programs and their impact on employee development. But we need information from people who managed the programs.

Attached is an evaluation questionnaire designed to depict your opinion related to our HRD programs and what type of information is more important from your perspective.
Role Playing Simulation

Organizational Role: HRD Trainer
Organization: Corporation of a Fortune 500 located in North America.
Subject: HRD Evaluation

Responsibility: Assess needs for HRD programs. Design and deliver training programs for all levels of the organization based on needs. For instance, you designed the supplemental retirement program and delivered a three hours hands-on computer training session for all sales personnel.

Rating Task: In 1992 the HRD department has presented a variety of programs for top management, middle management, and new employees. The purpose of the programs was to develop managerial skills and to improve efficiency.

As you know the company is cutting 30% of the budget for the human resources department next year. Therefore, the new philosophy is to evaluate our HRD programs without allowing the evaluation process cost to exceed potential benefits. Practical application of this philosophy ensures efficient evaluation utilization. The evaluation process is designed to: (1) Determine the impact of training programs on business objectives, and (2) Improve Corporate Training programs and their impact on employee development. But we need information from people who deliver the programs.

Attached is an evaluation questionnaire designed to depict your opinion related to our HRD programs and what type of information is more important from your perspective.
Role Playing Simulation

Organizational Role: Trainee
Organization: Corporation of a Fortune 500 located in North America.
Subject: HRD Evaluation

As a former participant in our Human Resource programs, and since you have had time to reflect on the program as a whole and attempted to put it into practice. We believe you would be a good source of advice for the evaluation of the program. We need your opinion regarding several aspects of the program. Rather than saying how well you liked the program you attended, please tell us how we should look at all programs to reflect the trainee point of view.

Training History: You have attended two different training sessions in the last year. Session one explained the new supplemental retirement deduction program. Session two was hands-on practice with a new computer software package in your department.

The new philosophy of our department is to evaluate our HRD programs without allowing the evaluation process cost to exceed potential benefits. Practical application of this philosophy ensures efficient evaluation utilization. The evaluation process is designed to: (1) Determine the impact of training programs on business objectives, and (2) Improve Corporate Training programs and their impact on employee development. But we need specific information from people who experienced a training module. Please take a few minutes to think about the entire program and give us your opinion. This information will be very helpful to us in planning future sessions.

Attached is an evaluation questionnaire designed to depict your opinion related to our HRD programs and what type of information is more important from your perspective.
Appendix E

Request for Subject Solicitation and Script
Dear <<Title>> <<LName>>:

My name is Fahad Al-Ruwaished. I am a Doctoral Candidate at Western Michigan University in the Educational Leadership program with an emphasis in Human Resource Development (HRD). I have chosen to research the importance of the evaluation questions and the information related to different clients and audiences.

I would like to ask your permission to include students from your class EDLD <<>> in my study. The students in your class will be assigned an HRD role (Trainees, HRD staff, HRD managers, or CEOs). They will then fill out a questionnaire about the importance of evaluation questions. I will brief them about this area of HRD evaluation if you are interested.

At no point will any individual be identified. Thus, the information they provide will be anonymous. I would appreciate it if you could give me twenty minutes from your class during the Winter Semester to conduct this study.

Sincerely,

Fahad Al-Ruwaished
Doctoral Candidate

Mary Anne Bunda, Ph.D.
Dissertation Chair
Dear Student:

My name is Fahad Al-Ruwaished. I am a Doctoral Candidate at Western Michigan University in the Educational Leadership program with an emphasis in Human Resources Development and my advisor is Dr. Mary Anne Bunda 7-3031. I have chosen to research the importance of the evaluation questions and the information related to different clients and audience. I would like to ask your permission to participate in my study.

At no point in your response will any individual be identified and you have the right not to participate. Thus, the information you provide will be anonymous.

Before distributing the questionnaire, the class will be divided into four groups. Each group will be assigned an HRD role (Trainee, HRD staff, HRD manager, or CEO). First, I would like you to read your simulation role. Second, you will fill out a questionnaire about the importance of evaluation questions according to your role. The task will take no longer than twenty minutes of your time. If you have any question, please feel free to ask.

Thank You
Appendix F

Human Subjects Institutional Review Board
Date: January 22, 1993

To: Fahad Al-Ruwaished

From: M. Michele Burnette, Chair

Re: HSIRB Project Number 93-01-19

This letter will serve as confirmation that your research protocol, "Human resources and evaluation: Clients and audience perspectives" has been approved under the exempt category of review by the HSIRB. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the approval application.

You must seek reapproval for any changes in this design. You must also seek reapproval if the project extends beyond the termination date.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: January 22, 1994

xc: Bunda, University Assessment


