

1981

The relationship between selected evaluator and decision maker characteristics and the impact of an evaluation

Valerie J. Ruffo Broughton
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Education Commons](#)

Recommended Citation

Broughton, Valerie J. Ruffo, "The relationship between selected evaluator and decision maker characteristics and the impact of an evaluation " (1981). *Retrospective Theses and Dissertations*. 7400.
<https://lib.dr.iastate.edu/rtd/7400>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB RD., ANN ARBOR, MI 48106

8209100

Broughton, Valerie J. Ruffo

**THE RELATIONSHIP BETWEEN SELECTED EVALUATOR AND DECISION
MAKER CHARACTERISTICS AND THE IMPACT OF AN EVALUATION**

Iowa State University

PH.D. 1981

**University
Microfilms
International**

300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark .

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Other _____

**University
Microfilms
International**

**The relationship between selected evaluator and decision maker
characteristics and the impact of an evaluation**

by

Valerie J. Ruffo Broughton

**A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY**

**Department: Professional Studies in Education
Major: Education (Research and Evaluation)**

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

**Iowa State University
Ames, Iowa**

1981

Copyright © Valerie J. Ruffo Broughton, 1981. All rights reserved.

TABLE OF CONTENTS

	<u>Page</u>
CHAPTER I. INTRODUCTION	1
Background	1
Need for the Study	2
Purpose of the Present Study	4
Hypotheses	7
Definitions	8
Limitations	9
CHAPTER II. REVIEW OF THE LITERATURE	11
Introduction	11
Related Research	11
Summary	26
CHAPTER III. METHODOLOGY	28
Subjects and Procedures	28
Materials	29
Data Analysis	37
CHAPTER IV. FINDINGS	40
Sample Profile	40
Reliability Estimates	42
Least Squares Regression Analysis	43
LISREL IV Analysis	53
CHAPTER V. SUMMARY AND CONCLUSIONS	61
Summary	61
Discussion	64a

	<u>Page</u>
REFERENCES	69
ACKNOWLEDGEMENTS	73
APPENDIX A. STANDARDIZED DIRECTIONS FOR RESEARCH MATERIALS	75
APPENDIX B. RESEARCH MATERIALS	78
APPENDIX C. PATH COEFFICIENTS AND RESIDUALS FOR ORIGINAL MODEL (FIGURE 4)	103
APPENDIX D. PATH COEFFICIENTS AND RESIDUALS FOR REVISION 1 (FIGURE 5)	107
APPENDIX E. PATH COEFFICIENTS AND RESIDUALS FOR REVISION 2 (FIGURE 6)	111
APPENDIX F. PATH COEFFICIENTS AND RESIDUALS FOR REVISION 3 (FIGURE 7)	115

LIST OF TABLES

	<u>Page</u>
Table 1. Sample Profile (Categorical Data)	41
Table 2. Sample Profile (Continuous Data)	42
Table 3. Summary Statistics for Measures	44
Table 4. The Relationship between Predictor Variables and Evaluator Quality	45
Table 5. The Relationship between Predictor Variables and Uses	46
Table 6. The Relationship between Predictor Variables and Usefulness of Process	47
Table 7. The Relationship between Predictor Variables and Usefulness of Information	48
Table 8. The Relationship between Predictor Variables and Overall Impact	50
Table 9. Correlation Coefficients	51
Table 10. Analysis of Variance: Usefulness of Process by Resume Treatment and Information Treatment	53
Table A. Residuals for Original Model	105
Table B. Residuals for Revision 1	109
Table C. Residuals for Revision 2	113
Table D. Residuals for Revision 3	117

LIST OF FIGURES

	<u>Page</u>
Figure 1. Evaluation utilization model	5
Figure 2. Evaluation utilization model	12
Figure 3. Evaluation utilization model	35
Figure 4. Original evaluation utilization measurement model	54
Figure 5. Revised measurement model (1)	58
Figure 6. Revised measurement model (2)	59
Figure 7. Revised measurement model (3)	60

CHAPTER I. INTRODUCTION

Background

The purpose of the present study was to identify factors which affect the usefulness of program evaluation information. During the last 20 years there has been an increase in the occurrence of evaluation efforts in schools. Throughout the 1960s, the War on Poverty hoped to cure the nation's ills by providing financial assistance to many social service agencies in order for them to implement new, and hopefully, effective programs. Schools were included in the population of agencies provided with funds and charged with improving societal conditions. The Congress of the United States desired evidence that the funds allocated to these agencies and schools were being used for the specified purpose and were accomplishing the desired ends. Thus, each project director was required to include an evaluation plan in every program proposal and to supply periodic evaluation reports to the funding body (Rossi et al., 1979).

Local school districts were anxious to take advantage of this national funding source and therefore complied with the legislative requirements. In addition, the accountability movement of the 1970s forced local districts to report to taxpayers concerning the effects of schooling. "Evaluation research was needed to provide data on how the many programmatic changes inflicted on the nation's children really made an educational difference" (Patton, 1978, p. 15). The emphasis on evidence and proof of a program's effectiveness resulted in a substantial increase in the occurrence of evaluations (Joint

Committee, 1981; Alkin, 1981). Hence, evidence collected during the last two decades indicates an increased demand for evaluation in schools.

The urgency of this growing demand found evaluators without well-developed guidelines. Therefore, evaluators had to improvise and borrow ideas from research methodologies. Problems with resulting evaluations have since been identified and solutions are currently being sought. Under-utilization of evaluation information is one such problem.

Need for the Study

Studying facets of evaluations that increase, impede, or neutrally affect the usefulness of evaluation information is one method of solving the utilization problem. The frequency of evaluations has recently increased. However, many authorities and experienced evaluators reported that the impact, or use, of evaluations was minimal (Braskamp & Brown, 1980; Locatis *et al.*, 1980). According to Worthen and Sanders (1973), "Evaluation is one of the most widely discussed but little used processes in today's educational system" (p. 1). Even when evaluation plans are implemented, they have little influence and much of their information is ignored (Guba, 1969; Mann, 1972; Weiss, 1972; Patton, 1978). The federal government has also shown concern about evaluations that are not used. In its 1979 Annual Report to the President and Congress, the National Advisory Council on the Education of Disadvantaged Children recommended that the Congress "...incorporate, within future legislative mandates for

national evaluation of the ESEA, Title I Program, specific purposes for the utilization of the information collected" (National Advisory Council, 1979, p. 2). Predicting that funds for evaluation would become unavailable if results did not become more useful, Atkisson (cited in Davis & Salasin, 1975) warned "...if evaluations do not result in more effective use, program managers will be reluctant to invest resources necessary to effect a workable evaluation strategy" (p. 625). Thus, the usefulness of evaluations became a focus of concern for educational researchers.

Both evaluators and consumers of evaluation recognized the problem of non- or under-utilization. Studies commissioned at the federal level supplied evidence that consumers of evaluations called for the study of the process of evaluation utilization. Recent federal initiative towards increasing the usability and utilization of evaluation findings included a large-scale NIE study conducted by Huron Institute. Also, the Educational Amendments of 1978 mandated that the Secretary of Education conduct a comprehensive review of federal evaluation practices and procedures (Boruch et al., 1981). One question investigated in this review process was: "How are the results of evaluation used?"

In addition, a body of research recently emerged in which evaluators studied conditions that maximize utilization (Brown et al., 1980; Braskamp et al., 1978; Dickey, 1980; Brown et al., 1978; Lorenzen & Braskamp, 1978; Locatis et al., 1980; Becker, 1981). This research began when one evaluator, Weiss (1972), proposed the need

for systematic empirical study of evaluation examining those conditions where evaluation is or is not used. Other evaluators also suggested the need to improve our understanding of utilization through empirical studies by examining factors that mediate effective utilization (Ciarlo, 1981; Conner, 1981; Braskamp & Brown, 1980). Thus, progress in the field of program evaluation requires investigation of specific evaluation conditions so that evaluators can be guided by methods which provide the most useful products.

Purpose of the Present Study

Based on the previous discussion that evaluators and consumers identified evaluation utilization as a problem and therefore that investigation of this problem is warranted, the present study was designed to determine whether certain decision maker and evaluator characteristics affected the usefulness of an evaluation. By developing a model based on ideas suggested by Alkin (1975), and varying evaluator and decision maker characteristics, the impact of a simulated evaluation report was measured. The researcher combined Alkin's suggestions with a parallel structure in communication theory to serve as the foundation for this study. In addition to verifying the model, this study tested relationships between background, experience, knowledge, and personality traits as predictors, and ratings of quality, uses, and usefulness of an evaluation as criteria. These findings produced implications for evaluators and decision makers, as well as suggestions for further research on this topic.

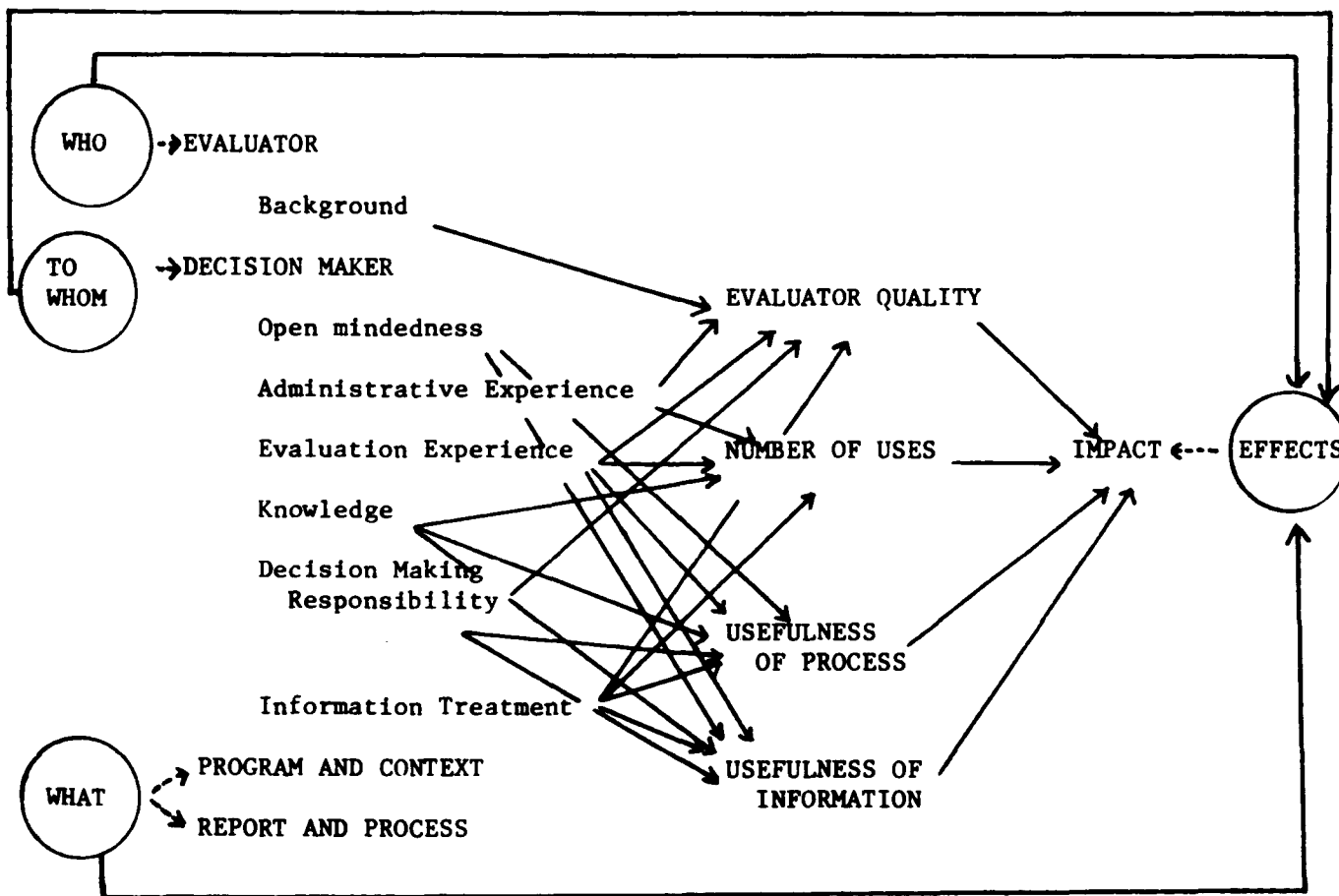


Figure 1. Evaluation utilization model

As can be seen, this model included four categories of factors which are believed to affect the usefulness of an evaluation: Evaluator Characteristics, Decision Maker Characteristics, Evaluation Report and Process, and Program and Social Context of an Evaluation. These categories correspond to a paradigm used in communication research: the effects of communication depend on who says what to whom (Triandis, 1971).

The present study investigated variables included in the evaluator and decision maker categories and held constant the program and context, and report and process factors. The evaluator variable studied was evaluator background. The six decision maker variables were: open mindedness, knowledge of program evaluation, administrative and evaluation experience, decision making responsibility, and an information treatment consisting of an essay about evaluation. The first level of criterion variables, which measured effects, were quality of the evaluator, uses of the evaluation, usefulness of the process, and usefulness of the information in the evaluation. These criteria contribute to the final measure, evaluation impact.

The population of decision makers who participated in the present study was practicing junior and senior high school principals or assistant principals. The Iowa State University Committee on the Use of Human Subjects in Research reviewed this project and concluded that the rights and welfare of the human subjects were adequately protected, that risks were outweighed by the potential benefits and expected value of the knowledge sought, that confidentiality of data

was assured, and that the informed consent was obtained by appropriate procedures.

Hypotheses

In this study, the evaluator variable, background, was operationalized by simulating resumes of two evaluators: one having a research background and the other having an evaluation background. The six variables included in the decision maker factor are operationalized as follows: Open mindedness will be measured using the Rokeach Dogmatism Scale (Rokeach, 1960). A 10-item quiz measuring general knowledge in the field of program evaluation developed by the researcher will be used to measure knowledge. The direct question, "For how many years have you been a school administrator?" measured administrative experience. Evaluation experience was measured by asking respondents to indicate in which of 18 evaluation situations they had participated. Decision making responsibility was measured by combining the variables, position, size of district, size of school, and influence the administrator had on personnel decisions. The information treatment consisted of an essay comparing and contrasting educational research and educational evaluation adapted by the researcher from Popham (1975).

The arrows in Figure 1 depict the hypothesized relationships tested, which were:

1. There is no linear relationship between evaluator background, administrative experience, evaluation experience, decision making responsibility, and the information treatment as predictors of the criterion, evaluator quality.

2. There is no linear relationship between knowledge, administrative experience, evaluation experience, and the information treatment as predictors of the criterion, uses.
3. There is no linear relationship between open mindedness, knowledge, evaluation experience, decision making responsibility, and the information treatment as predictors of the criterion, usefulness of the evaluation information.
4. There is no linear relationship between open mindedness, knowledge, evaluation experience, decision making responsibility, and the information treatment as predictors of the criterion, usefulness of the evaluation process.
5. There is no linear relationship between evaluator quality, uses, usefulness of process, and usefulness of information as predictors of the criterion, overall impact.

In addition to testing these specific relationships, the evaluation utilization model was tested and revised.

Definitions

At this point it is necessary to define terms which may be unfamiliar or which may have confusing connotations. The starred (*) definitions were taken directly from the Standards for Evaluations of Educational Programs, Projects, and Materials, one purpose of which was "to provide a set of working definitions to guide research and development on the evaluation process" (Joint Committee, 1981, p. 5).

Alternative utilization -- Use which provides an understanding of a program and facilitates discussions without directly influencing a specific decision. Indicates an indirect influence of evaluation information.

*Context of an evaluation -- The combination of the factors accompanying the study that may have influenced its results. These factors include the geographic location of the study, its timing, the political and social climate in the region at that time, the other relevant professional activities that were in progress, and any existing pertinent economic conditions.

Evaluation -- Systematic investigation of the worth or merit of an object; e.g., a program, a project, or instructional material, for the purpose of influencing decision processes either directly or indirectly.

***Evaluator -- Anyone who accepts and executes responsibility for planning, conducting, and reporting evaluations.**

***Executive summary -- A summary statement designed to provide a quick overview of the full-length report on which it is based.**

Mainstream utilization -- Use in which an evaluation has immediate and direct influence on a program, project, or material which results in immediate change or innovation.

Joint Dissemination Review Panel -- A federal review board that examines educational products and practices and determines whether the submitting project has provided persuasive evidence of effectiveness. Programs that pass JDRP's screening procedure are eligible for dissemination funds from the National Dissemination Network.

***Program evaluations -- Evaluations that assess educational activities that provide services on a continuing basis and often involve curricular offerings. Examples include evaluation of a school district's reading program, a state's special education program, and a university's continuing education program.**

***Project evaluations -- Evaluations that assess activities that are funded for a defined period of time to perform a specified task. Some examples are a three-day workshop on behavioral objectives, a two-year development effort, or a three-year career education demonstration.**

Simulation study -- A study that uses symbolic representation of real activities, situations, or environments.

***Utility -- The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work.**

Limitations

This study shared a limitation with all simulation studies; it has limited generalizability. Because of the specificity of the contrived situation, it was difficult to generalize directly to the

real world. However, when the present findings were considered along with other evaluation utilization research findings, a meaningful discussion resulted.

Organization

Subsequent chapters of this dissertation include a report of current evaluation literature as it relates to this study, descriptions of the methods and procedures, findings of the study, a discussion of the relationship between these findings and other relevant research findings, and implications for evaluators and decision makers.

CHAPTER II. REVIEW OF THE LITERATURE

Introduction

The purpose of the present study was to test a model consisting of four factors hypothesized to determine the usefulness of an evaluation. The factors in the model were suggested and supported by Alkin (1975). This model combines a communications theory paradigm (Triandis, 1971) with the outline presented by Alkin. The four factors included in the utilization model were: evaluator, decision maker, context and program, and process and report. Only the first two factors were investigated in the present study. The variables operationalizing these two factors measured the effect of evaluator and decision maker characteristics on the impact of a simulated evaluation study. As can be seen in Figure 2, the seven variables studied were: evaluator background, decision maker open mindedness, knowledge of program evaluation, administrative and evaluation experience, decision making responsibility, and information about evaluation.

Related Research

In this chapter, the author will review related literature in the field in evaluation. Each of the four factors in the model will be described and supported. All of the works cited will relate to one of the factors defining the evaluation utilization model, as shown in Figure 2.

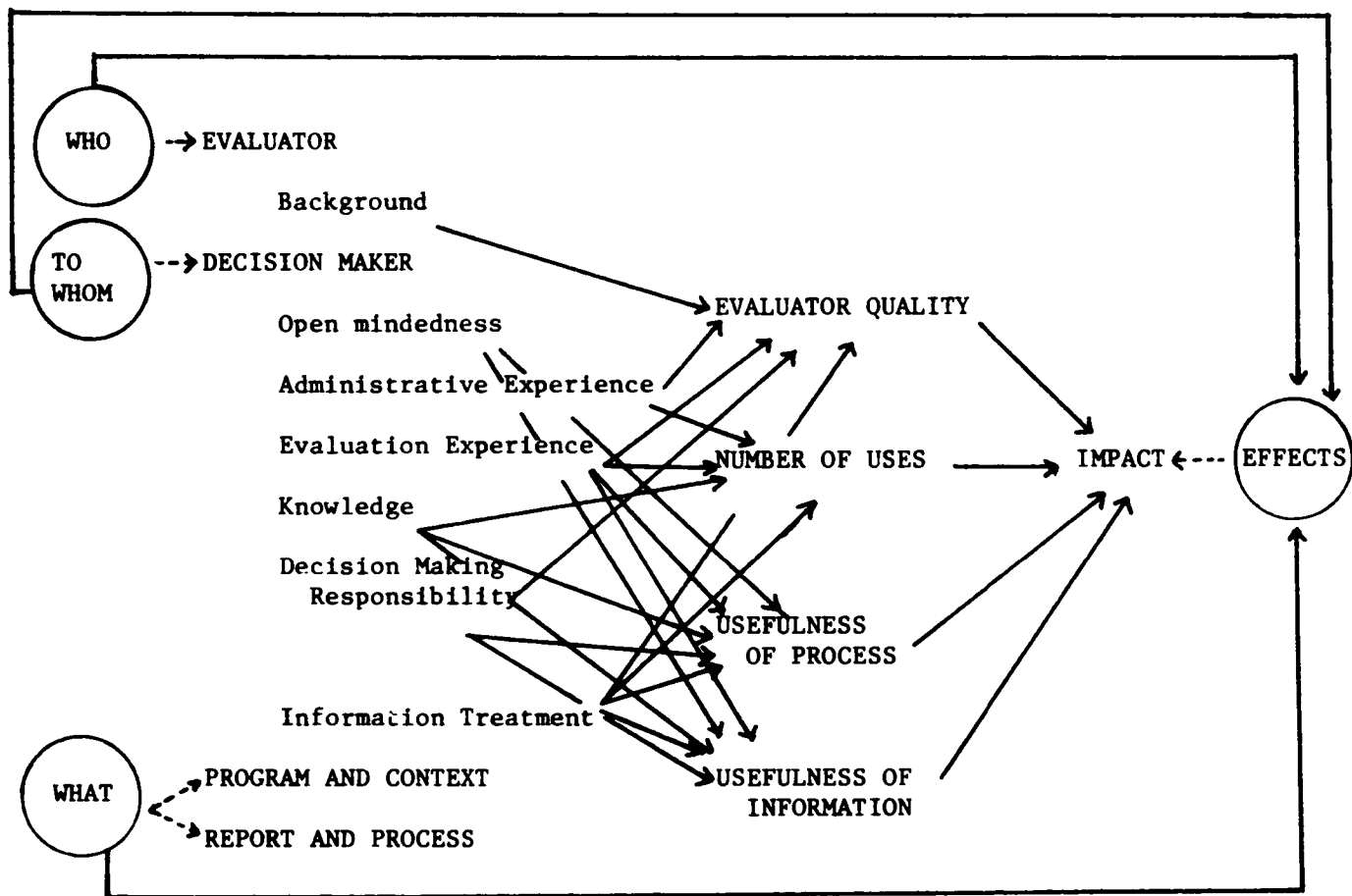


Figure 2. Evaluation utilization model

Recommendations of researchers from previous studies

When it became apparent that the impact of many evaluation studies was minimal, evaluators called for the study of the evaluation process and how that process contributed to the ultimate usefulness of an evaluation report. Originally, it was assumed that characteristics of the evaluator and the report which assured quality also assured usefulness. Therefore, many lists of suggestions designed to insure the quality of an evaluation report were developed (Cronbach, 1963; Phi Delta Kappa, 1971; Weiss, 1972; Davis & Salasin, 1975; Braskamp & Brown, 1980; Anderson & Ball, 1978; Haenn & Owens, 1981).

The five recommendations mentioned most often in the above speculative discussions were related to the evaluator, decision maker, and report and process factors in the model.

1. **Timeliness** -- The evaluation report must be completed while the program is still malleable.
2. **Purpose and involvement** -- The evaluator must know the needs of the decision maker and be aware of his/her values. The most frequently suggested method for accomplishing these is to first identify the decision makers and then involve them in all phases of the evaluation.
3. **Evaluation questions** -- The questions studied by the evaluator should be specific, narrowly focused, and endorsed by the decision maker.
4. **Dissemination** -- Informal reports should be made continuously to the program staff and decision makers. The final reports should be tailored to each specific audience, its interests and needs.
5. **Assistance** -- The evaluator should be helpful after the report is complete by suggesting means by which the recommendations can be adopted.

The description of purpose and involvement suggests a relationship between the decision maker factor and the usefulness of an evaluation. The last recommendation, assistance, defines a role of the evaluator. Timeliness, evaluation questions, and dissemination are variables contained in the report and process factor of the model.

The significance of the evaluation process and the evaluation report was highlighted again in a recent publication which outlined suggestions of how to successfully carry out an evaluation, Standards for Evaluations of Educational Programs, Projects, and Materials, written by the Joint Committee on Standards for Educational Evaluation (1981). The committee categorized a set of comprehensive standards by the four most important attributes of an evaluation: utility, feasibility, propriety, and accuracy. Utility is the factor most related to the present study and contains standards for guiding evaluations so that they will be informative, timely, and influential. Thus, the Joint Committee also recognized these report and process variables.

Ethnographic studies

Evaluators studying the problem of under- or non-utilization recently completed two studies conducted specifically to examine utilization in actual field settings (Patton, 1978; Alkin et al., 1979). Results of both studies include descriptions of evaluations that were classified as useful, as well as the characteristics of the evaluations which contributed to their usefulness. All four factors

included in the evaluation utilization model are supported by findings of these two studies and are discussed below.

Participants in an evaluation methodology training program at the University of Minnesota decided to find a few examples of evaluations that were utilized and study those examples to learn how to increase utilization. The goal of this project was to develop a "comprehensive approach to program assessment" that provides a framework within which planners can develop an evaluation design with a built-in utilization component (Patton, 1978). This group conducted follow-up studies of 20 federal health evaluations. They attempted to assess the degree to which these evaluations had been used and to identify factors that affected varying degrees of utilization. They interviewed three people for each of the 20 evaluations: the project officer, the person identified by the project officer as the decision maker for the program, and the evaluator. Examining these evaluations in detail, the investigators found that political considerations were important factors explaining utilization. These considerations are associated with the program and context factor included in the evaluation utilization model.

In addition to identifying the importance of political considerations, Patton suggested that there are two fundamental requirements for producing a useful evaluation. These two requirements are related to the evaluator and decision maker factors investigated in the present study. First, relevant decision makers and information users must be identified and organized. Second, evaluators must work

actively, reactively, and adaptively with these identified decision makers and information users to make all decisions about the evaluation. Patton designated these as the "Personal Factor." The components of the Personal Factor reflect the evaluator and decision maker factors in the model investigated in the present study.

Stressing the interaction between evaluator and decision maker, Patton outlined these four additional suggestions for improving the usefulness of evaluation results.

1. The relevant evaluation questions must be identified and focused.
2. Selected evaluation methods should generate useful information for decision makers.
3. Decision makers and information users must participate with evaluators in data analysis and data interpretation.
4. Evaluators and decision makers must negotiate and cooperate in dissemination efforts.

Thus, the importance of the evaluator and decision maker facets of the evaluation utilization model is further supported.

One additional suggestion resulting from Patton's research was that a workable definition of utilization was needed. Although this discussion is not directly related to the four factors in the utilization model, it is related to the measurement of the utilization variables. One significant finding was that 78 percent of responding decision makers and 90 percent of responding evaluators felt that the evaluation had an impact on the program. This is contrary to the concerns of many evaluators that evaluations are not used (Guba, 1969; Weiss, 1972; Ciarlo, 1981; Conner, 1981; Braskamp & Brown, 1980; Boruch et al., 1981). However, the impact found by Patton was not

something that suddenly and concretely occurred at one specific moment. Rather, it was the result of a gradual process which reduced decision maker uncertainty. Therefore, Patton recommended broadening the narrow definition commonly used when considering the usefulness of an evaluation. The present study examined many kinds of potential uses in an attempt to capture the meaning of the broad definition of evaluation utilization.

In addition, another field study suggested the same definitional clarification. Alkin et al. (1979) proposed two types of evaluation utilization: mainstream and alternative. The mainstream definition uses immediate and direct impact on one or more critical program decisions as the criterion for utilization. The alternative definition of evaluation utilization focuses less on the notion of dramatic impact on program decisions and includes gradual influence on administrator perceptions which may slowly change the course of the program. The alternative definition allows for the fact that evaluation is only one of the many inputs into a decision. In addition to the immediate versus gradual distinction between mainstream and alternative utilization, Alkin et al. contrasted them on another point.

The mainstream literature concentrates heavily on fairly static factors influencing (usually impeding) utilization. ...the alternative perspective on utilization emphasizes the interactions of people and situations in the evaluation process rather than looking simply at single factors or even at configurations of (fundamentally static) factors (p. 25).

In order to capture the complex nature of utilization, the uses measured in the present study represent the broad or alternative definition.

The results reported by Alkin et al. (1979) also included a tentative framework of factors which should maximize both types of utilization and which are related to the evaluation utilization model in the present study. Alkin et al. (1979) developed case studies that focused on ESEA Title I or Title IV-C programs and described a complete and accurate picture of the evaluation process. The researchers examined "the persons who shaped the process, how the evaluation fit into the total operation of the school program, and in what way the evaluation influenced decisions made about the program" (p. 35). Following are the eight categories in Alkin et al.'s analytic framework: preexisting evaluation bounds, orientation of the users, evaluator's approach, evaluator credibility, organizational factors, extra-organizational factors, information content and reporting, and administrator style.

The categories of preexisting evaluation bounds, organizational factors, and extraorganizational factors define the program and context factor contained in the evaluation utilization model investigated in this study. The seventh category, information content and reporting, reflects the report and process facet of the model. The decision maker factor is highlighted by the orientation and administrator style components of this framework. The evaluator, or who, factor in the evaluation utilization model includes evaluator approach and credibility.

Thus far, recommendations of evaluators and schemes resulting from field research have been presented. The last section of this

chapter includes specific studies supporting each of the four factors in the evaluation utilization model.

Evaluator

Support for the first factor of the model follows. Archibald (1970) confirmed the importance of the evaluator's reputation and legitimacy as strongly related to the potential utilization of evaluation findings. Another study which examined the relationship between evaluator reputation and credibility and the impact of an evaluation (Braskamp et al., 1978) found that an educational researcher was rated higher in objectivity than either an evaluation specialist or an art educator. However, these titles had no effect on respondents' perceptions of the usefulness of the evaluation information.

The present study will replicate the procedure of Braskamp et al. by varying the evaluator factor. One-half of the decision makers will read a resume of an evaluator having a "research background" and the other half will read a resume of an evaluator having an "evaluation background."

Decision maker

Alkin (1975) stated that improper recognition of who makes what decisions is a major impediment to utilization of evaluation findings. Also, even after decision maker identification, the evaluator must realize that the value system, philosophical orientation, political ambitions, and personality makeup, and "who knows what else" affect the extent to which evaluation findings are utilized. The decision purpose must also be considered in order to maximize the usefulness of an evaluation.

"The surest predictor of likely utilization is still an intended utilization at the outset of the project" (Alkin, 1975, p. 201).

Other findings related to this category and based on interviews with program decision makers follow. An evaluation activity initiated by a decision maker is more likely to be used than an evaluation initiated under any other circumstances (Horowitz, 1981). Evaluations that had great impact were planned considering the decision making orientation of the chief administrator in the school (Stecher, 1981). Davis (1981) found that the management style of the decision maker affected subsequent utilization of the evaluation. Elementary principals who were "compliance" oriented, i.e., saw their purpose as related to the policies and procedures of the board of education, found standardized test data information most useful. On the other hand, "non-compliance" oriented elementary principals, i.e., who saw their role as coordinating staff and encouraging high morale, developing programs, and planning, looked to data generated within the school as most helpful for decision making.

Also, in a series of studies conducted at the Huron Institute, relationships between decision makers and utilization of evaluation information were investigated. Principals (Neumann, 1981) and program managers (Apling, 1981) said that evaluations which included mention of ways the information could be considered were most useful to them. Kennedy (1981) discovered that policy makers use descriptive data and that in a group decision making situation, causal inferences based on group members' experience were linked to the data to facilitate decisions.

The following findings related to the decision maker factor were based on experimental studies. Decision maker attitude (Dickey, 1980), perceived need for evaluation (Brown et al., 1980), and position (Braskamp et al., 1978) are decision maker characteristics found to affect the usefulness of evaluation information. Conflicting findings are reported by Brown et al. (1980) concerning decision maker information about evaluation. In one study, the group of subjects who read an article about the importance of evaluation rated the usefulness variables higher than the group who did not read the article. However, in a similar study, they found no difference based on the article treatment. Years in current administrative position (Lorenzen & Braskamp, 1978) and knowledge of the object of the evaluation (Locatis et al., 1980) had no effect on decision maker judgments about simulated evaluation reports.

The present study investigated six decision maker characteristics to determine their relationship to evaluation utilization. Because evaluation utilization was found to depend on variables such as decision making orientation (Stecher, 1981) and management style (Davis, 1981) in two personal interview studies, the researcher in the present study sought to test a similar variable empirically. Dogmatism, or closed mindedness of decision makers, was the variable selected. Specific support for studying the effect of dogmatic orientation of decision makers is:

Research using the Rokeach Dogmatism Scale (Rokeach, 1960) supports the view that the dogmatic decision maker (i.e., one with a closed belief system) is characterized by rapid decisions based upon relatively little information, yet

once made, those decisions are confidently and inflexibly held (e.g., Block & Peterson, 1955; Brenglemann, 1975). Accordingly, dogmatic has been interpreted by Long & Ziller (1965) as a defense mechanism that inhibits predecisional information processing. This conclusion has received further support from the positive association found between dogmatism and anxiety (Rokeach, 1960). Limited pre-decisional information search was hypothesized to serve as a defense; it closes the mind to new information and eliminates any need for the decision maker to reevaluate his self concept. The dogmatic decision maker would, therefore, be expected to unduly restrict his information input. The resulting reduction in his information capacity would severely handicap his ability to apply strategies for coping with decision problems, ... (MacCrimmon & Taylor, 1976, p. 1439).

The main purpose of the dogmatism scale is to measure individual differences in open and closed belief systems. In the present study, it was hypothesized that this construct affected decision makers' perceptions of the usefulness of evaluative information.

Two other variables examined in the present study relate to findings presented in this decision maker section: knowledge and information treatment. One-half of the decision makers involved in the present study read an essay describing the difference between research and evaluation; the other half did not. The effect of this treatment will add to the findings reported by Brown et al. (1980). In addition, a 10-item quiz measuring general knowledge of the field of educational program evaluation was administered to all subjects.

The years of administrative experience variable in this study will replicate the findings of Lorenzen and Braskamp (1978). In addition, the effect of a more relevant variable, evaluation experience, will be investigated. Position of respondents could not be used to predict subjects' decisions or rating of usefulness (Braskamp et al.,

1978). However, a more comprehensive variable, decision making responsibility in current position, might. In this study, a decision making responsibility scale was constructed by combining size of district, enrollment in building, influence of personnel decision input, and position. All subjects were either principals or assistant principals at the secondary level.

Evaluation process/evaluation report

Several studies which investigated the relationship between report style and/or evaluation process and the usefulness of an evaluation will be presented.

Alkin (1975) described six characteristics of the evaluation report and the process used during an evaluation study that should assure its usefulness:

1. Attention to appropriate goals.
2. Technical credibility.
3. Report comprehensibility.
4. Report timeliness.
5. Scope of recommendations.
6. Evaluator relationships.

In a project commissioned by the Department of Education, Boruch et al. (1981) studied federal level evaluation reports and interviewed those who participated in the process. Among their results were recommendations to the Department of Education suggesting methods of conducting and reporting evaluations so as to increase the utility of evaluations. Some of the recommendations were, that the evaluation staff:

1. Provide oral reports regularly as well as written reports on results of major evaluations, and on the uses to which reports can be put.
2. Create a system to periodically collect, synthesize, and report specific uses to which evaluations are put.
3. Direct evaluation staff to meet regularly with congressional staff to clarify information needs, feasibility of evaluation, audiences for results, and ways in which results can be used to modify programs.

These suggestions could be applied to any evaluator who was interested in improving his/her evaluation techniques.

Another study which examined completed evaluation reports, and questioned involved personnel, was conducted by Dickey (1980). She studied 47 Title IV-C program evaluations in Minnesota and the decision makers associated with each program. Finding that half of the decision makers rated their evaluation as useful to very useful, Dickey tested the relationship between several variables and the evaluation's usefulness. She found timely completion of the evaluation, and data collection procedures related to usefulness. The report format, whether or not the report included recommendations, and decision maker involvement were not related to usefulness.

In a simulation study, Brown et al. (1980) conducted a three-part investigation of evaluation utilization and produced findings that linked report and process variables to utilization. They found that a report which provided data to support the recommendations was rated as having better quality and quantity of information than a report which included only recommendations and no data. The decision makers also were more satisfied with the data supported report.

In a similar study, Brown et al. (1978) investigated the effect of report style on decision maker rating of the technicality, difficulty, and acceptance of the evaluation. They found that an objective report containing educational jargon was rated as more difficult than reports which were subjective and/or jargon-free. However, report acceptance did not depend on report style.

In another study investigating the process and report factor, Lorenzen and Braskamp (1978) examined the influence of three types of evaluation information on administrative decision making. The three types of information contained in a simulated evaluation report were political, cost/benefit, and statistical information. Cost/benefit information was found to be most useful.

The effect of presentation style on audience reaction to a program evaluation report was studied by Becker (1981). Teachers and aids rated informal reports higher than formal reports on readability, attractiveness, presentation of numerical information, report summary, comprehensibility, interestingness, objectivity, relevance, credibility, and usefulness.

The last research study investigating the third factor in the evaluation utilization model, report and process, revealed that the kind of information included in the report -- positive, negative, neutral, or conflicting -- affected decision maker judgments of quality (Locatis et al., 1980). They found that evaluative information that is not uniformly positive tends to lower ratings.

In the present study, the report and process factor was held constant. Considering the research findings presented above, the author developed the simulated report and process to involve the program participants in all phases of the process, to frame specific evaluation questions, to include data to support the recommendations, to be objective and use educational jargon sparingly, to consider both the cost of the program and of implementing the recommendations, to follow a functional rather than a formal report format, and to complete the report in a timely major.

Program and social context

Few research studies related to the fourth factor of the model, program and social context, could be located. The nature of the program, characteristics of the program staff, formal and informal organizational structure, and political alignments of the various external groups were decisive determiners of utilization/non-utilization (Deats, 1974; Ferman, 1969). In addition, Dickey (1980) found that evaluations of programs which had been validated by the Joint Dissemination Review Panel were rated higher in usefulness than evaluations of programs not validated. The program and context for the present simulation study were held constant and therefore not investigated.

Summary

The two factors thought to predict evaluation usefulness investigated in the present study were evaluator and decision maker

characteristics. The personal factor (Patton, 1978) stressed the importance of the interaction between the evaluator and decision maker in producing useful evaluation results. Currently many people charged with evaluating programs were trained in other fields. Therefore, it is necessary to study the effect of evaluator background on evaluation usefulness. Alkin (1975) proposed the importance of studying the relationship between decision maker characteristics and utilization. Information presented in this chapter indicated that decision maker and evaluator characteristics do affect the use of evaluation information.

The researcher also provided information concerning the two factors in the model that were not investigated in this study. The findings related to the program and context factor of the model pointed to the importance of the evaluator's awareness of any political situations surrounding a program which might interfere with the decision maker's using the evaluation results. Several studies investigating the process and report factor were reported in this chapter. Most reinforced what an evaluator's common sense direct him/her to do: complete the project on time, write the report so that it is readable to the audience, include supporting data in the report, and consider the cost of the recommendations.

The details of how the present study was conducted will be discussed in the next chapter.

CHAPTER III. METHODOLOGY

Subjects and Procedures

The population of subjects participating in this study was practicing secondary school administrators. Specifically, this group included high school and junior high school principals or assistant principals. A total of 80 administrators cooperated. Of the total, 3 were female, 77 were male; 35 were principals. The range of years of administrative experience for participants was 1 to 29.

The researcher personally administered the treatments to 10 groups of principals. The groups ranged in size from 3 to 17. All groups met for one and one-half hours during the summer of 1981. The researcher standardized directions and administration procedures to minimize condition effects on responses (see Appendix A).

The researcher contacted central office administrators in 12 of the 20 districts in Iowa with the largest enrollments and asked them to cooperate in this research study; nine agreed. The contact person recruited volunteers from his district and arranged a meeting at which the researcher conducted the study. The researcher traveled to all nine communities to meet with groups of principals. Each meeting was held in a conference room of the central office building or in a high school classroom. Sixty-six subjects were contacted in this manner.

The remaining 14 subjects were enrolled in a Supervision and Evaluation workshop conducted at Iowa State University in July, 1981. This group completed the research materials during the last meeting of

the workshop. The majority of these principals was from Iowa. However, several other states were represented (New York, Illinois, Georgia). They met in the back of the workshop classroom. While the researcher administered the materials to the group in the back of the classroom, the participants in the workshop who were teachers, and therefore not participants in the research, were listening to a lecture in the front of the same classroom. Therefore, although the materials and treatment administration were the same for the workshop subsample as for the district subsample, there were definite distractions for the workshop people.

Materials

The purpose of this study was to measure the effect of evaluator and decision maker characteristics on the impact or overall usefulness of a simulated evaluation. Evaluator and decision maker characteristics are two of the four major categories in the evaluation utilization model previously discussed.

The materials used in this study consisted of a packet containing three sections. First, there was a questionnaire section asking for information about decision maker characteristics: open mindedness, knowledge, administrative and evaluation experience, and decision making responsibility. The second major section included the simulated evaluation materials: the information treatment, a description of the program and context, a resume of the evaluator, a letter, and the

evaluation report. The report was an executive summary which included recommendations and was organized according to four questions investigated by the evaluator. The last section of the materials packet contained questions measuring decision maker perceptions about the criteria: quality of the evaluator, uses of the evaluation, usefulness of the information, and usefulness of the process of the four questions discussed in the report, and finally, the overall usefulness of the evaluation, or impact. The pages were color coded for ease of reference when administering the materials (see Appendix B for all materials used).

Decision maker characteristics

Both Alkin et al. (1979) with their framework, and Patton (1978) with his personal factor, stress the contribution of decision maker characteristics to the ultimate usefulness or impact of an evaluation. Following are descriptions of the specific decision maker characteristics investigated. The first is open mindedness which refers to a score on the Rokeach Dogmatism Scale Form E (Rokeach, 1960). This scale consists of 40 items, has the stated purpose of measuring individual differences in open and closed belief systems, and has a reported reliability of .68-.93 (Rokeach, 1960, p. 89).

The second decision maker characteristic investigated in this study was the administrator's knowledge of program evaluation. The 10-item quiz measuring this variable was developed by the researcher and pilot tested on a group of 37 education administration graduate students. The range of scores on this pilot was 0-8, and coefficient

alpha was .60. These statistics were computed again for the actual sample.

Next, subjects were asked to report the number of years they had been a school administrator as a measure of the third decision maker characteristic, years of administrative experience. Subjects were directed to round off any partial years to the nearest whole year. It was assumed that the response would be truthful and accurate.

The next decision maker variable, evaluation experience, was operationalized by asking administrators to check whether they had experienced certain evaluative activities. The researcher generated a list of 18 activities. Three of the items asked if the administrator was "directly responsible" for an evaluation activity. The other 15 asked if the administrator was "involved" in an evaluation activity. Responses to the three "directly responsible" items were weighted doubly. A reliability estimate for this list was calculated to determine whether only one construct was being measured by this scale, or whether more than one factor was represented.

The last decision maker characteristic investigated in this study was decision making responsibility. This was obtained by standardizing and combining responses to questions on position, size of district, size of school, and the influence administrators' opinions had on personnel decisions at their school. It was assumed that principals would have more decision making responsibility than assistant principals. Further, it was assumed that administrators from larger schools in larger districts would have more decision making responsibility than

those from smaller buildings and districts. Also, a positive relationship was assumed to exist between the impact that an administrator's opinion had on personnel decisions at his/her school and the administrator's decision making responsibility.

Simulated evaluation materials

This second section of materials included descriptions of the program and context of the evaluation and the evaluation process and report, as well as the information and evaluator treatments.

Treatments In addition to these decision maker characteristics, there were two treatment effects investigated in this study: the resume treatment (an evaluator variable) and the information treatment (a decision maker variable). These treatment materials were included in the second section of the packet which contained the simulated evaluation. One-half of the subjects received a packet containing a resume describing the evaluator as a public school central office administrator with experience as a teacher, principal, and evaluator, and no publications in the field of evaluation. This resume treatment is described as the evaluator having an "evaluation background." The other half of the participants read a resume of an evaluator with a "research background." This person had identical training and public school teaching experience as the first evaluator. However, he had no direct experience as an evaluator, but had published several research articles about program evaluation. The present study assessed the effect of the background of an evaluator on decision makers' perceptions of the quality of the evaluator, and on the impact of the evaluation.

The information treatment was included in this study to determine whether an administrator's awareness of the difference between research and evaluation affected the impact of the evaluation. The information treatment consisted of a one-page essay titled, "A Comparison and Contrast of Educational Research and Educational Evaluation." Adapted by the researcher from Popham (1975), it included a discussion of the differences between research and evaluation as well as their similarities. The validity of this treatment was established by asking a group of 14 education graduate students to read the essay and then answer factual questions about the content to determine if they gained the knowledge the author had intended. Eighty-one percent answered all four questions correctly. Therefore, it was decided that the essay would accomplish its purpose of teaching the readers the difference between research and evaluation.

Therefore, summarizing the organization of these two treatments: 20 subjects read the essay and an "evaluation background" resume (coded 11); another 20 read the essay and a "research background" resume (coded 12); 20 more did not read the essay but read the "evaluation background" resume (coded 01); and the last 20 did not read the essay but read the "research background" resume (coded 02).

Program and context Other materials included in the evaluation simulation section of the packet reflect the two components of the evaluation utilization model not investigated in this study: the report and process, and the program and context. The materials describing these components were the same for all 80 subjects. The program being evaluated in this simulation was a two-year-old remedial math course still in its formative stages. The context was such that

the teachers had requested that an outside evaluator help them improve their program. The descriptions of the program and context were each one page in length.

Report and process A description of the evaluation process and the evaluation report were included in the last six pages of the second section, simulated evaluation materials. The math teachers were to have selected four questions they wanted the evaluator to study in depth. The researcher generated a list of 19 possible questions and asked a group of 20 education administration graduate students to select the four most important. The four selected most often were used in the simulated report. The simulated report and process were written to include the components of a successful evaluation discussed previously: timeliness, inclusion of recommendations, client involvement, and clearly stated evaluation questions.

Criterion measures

The third section of the packet administered to the subjects consisted of three pages of questions measuring the usefulness of the simulated evaluation. As can be seen in Figure 3, there are two levels of dependent variables. The first level includes measures of the quality of the evaluator, the number of uses of the evaluation, and the usefulness of the process and usefulness of the information of the four evaluation questions investigated by the evaluator. The variable at the second level is the impact or overall usefulness of the evaluation.

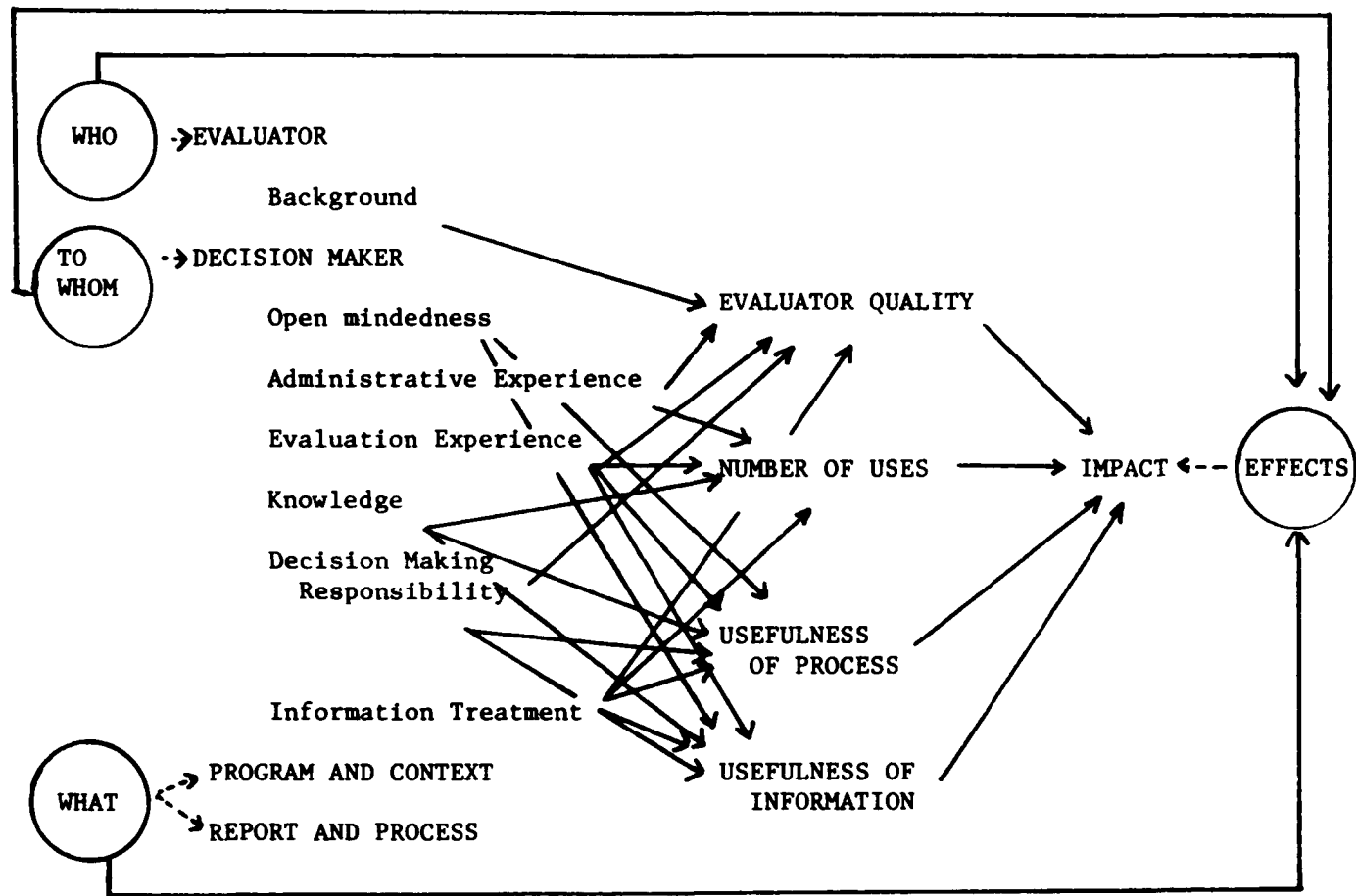


Figure 3. Evaluation utilization model

Quality of the evaluator was measured by asking decision makers to rate the evaluator on 12 traits. Also, a question asking whether the decision maker would recommend the evaluator to a respected colleague assessed the quality of the evaluator. The reliability of the scale which included all 13 of these variables was estimated.

The second criterion included in the model is number of uses of the evaluation. It was assessed by asking decision makers which of nine possible uses were applicable to the given simulation setting. Subjects were instructed to check all that applied. The number of checked responses was totaled. The author of the present study sought a checklist that would include choices describing indirect or delayed impact, i.e., alternative utilization (Alkin et al., 1979). The list used in this study was developed with that purpose in mind (Dickey, 1980).

The next two measures in the first level of dependent variables are similar. One scale, the process scale, was devised by adding responses from the question, "How would you rate the usefulness of the evaluation process?" for each of the four evaluation questions investigated by the evaluator. The other scale, the information scale, totaled responses to "How would you rate the usefulness of the information?" for each of the four evaluation questions. This technique is an adaptation of a measure used by Dickey (1980). The use of these two scales also reflects the broad definition of utilization measured in the present study. Thus, the four variables measuring the first level of usefulness are quality of the evaluator,

uses of the evaluation, and usefulness of the evaluation process and usefulness of the evaluation information.

The second level consisted of one scale which included three questions. Considering the evaluation in total, the decision makers were asked to rate the usefulness of the information, process, and total evaluation. These three responses were totaled to form a measure of overall impact, or utilization. Dickey (1980) used these three questions in her study. An estimate of the reliability of the scale used in the present study was calculated.

Data Analysis

The author coded the packets and the keypunch section of the computation center prepared the data cards. Descriptive, reliability, and preliminary regression statistics were computed using the Statistical Package for the Social Sciences (Nie et al., 1971). The .05 level of significance was used to judge the significance of the regression coefficients. Summary statistics describing the sample are provided in the next chapter. In addition, reliability findings for each scale described previously are reported.

The evaluation utilization model was finally tested using the LISREL IV computer package (Joreskog & Sorbom, 1978) permitting a path analysis interpretation which adjusts for measurement error and examines structural equations. It is based on Joreskog's development of a maximum likelihood solution for linear structural equation systems. Because many readers may be unfamiliar with the LISREL approach to analyzing structural equations, three unique advantages of this

technique are discussed. They are measurement error, multiple indicators, and a chi-square test for overall goodness of fit.

The LISREL path analysis technique is based on multiple regression. In regression, the assumptions of homoscedasticity, normality of disturbances, and interval level of measurement can be violated without too much effect. However, measurement error can cause undesirable fluctuations in regression coefficients (Bohrnstedt & Carter, 1971), and can also lead to faulty inferences in path analysis (Blalock, 1965). The researcher knew of two methods of accounting for problematic measurement error.

Of these the errors-in-variables approach of Warren, White and Fuller (1974) and Joreskog's use of the covariance structures in LISREL and in earlier more specific cases (e.g., 1970), have been particularly useful (Evers, 1979, p. 153).

Therefore, the coefficients estimated by the LISREL technique varied from the ordinary least squares coefficients because the former were corrected for measurement error.

The use of multiple indicators is another major advantage of the LISREL program. Factor analysis and multitrait-multimethod analysis (Campbell & Fiske, 1959) are examples of other techniques employing multiple indicators. In ordinary least squares, a composite of items must be formed and entered into the equation as one predictor. If the items are measured on different scales, they usually need to be standardized before they are combined. The LISREL approach considers each item in a composite, and weights each item appropriately.

The third important benefit of the LISREL approach is the overall test of the fit between the data and the proposed model. "The goodness of fit is tested between the model and the data by comparing the variance-covariance (dispersion) matrix estimated by LISREL with the observed dispersion matrix. The null hypothesis is that any deviations between the estimated and observed matrices are due to chance" (Evers, 1979, p. 157). Therefore, a statistically significant chi-square value indicates rejection of the null hypothesis leading to the conclusion that the model does not fit the data. A non-significant chi-square suggests that the data do fit the model.

Because of the advantages offered by this technique, the LISREL method was used to supplement the ordinary regression results and to test and refine the evaluation utilization model.

CHAPTER IV. FINDINGS

In the present study, the researcher examined the relationship between evaluator and decision maker characteristics and tested their ability to predict decision makers' perceptions of the usefulness of evaluative information. Both ordinary least squares regression and path analysis techniques were used to investigate the relationships. The path analysis technique required that a reliability estimate be obtained for each criterion and predictor variable. In addition, descriptive statistics were calculated to summarize the background of the subjects, secondary school administrators.

Sample Profile

The population of decision makers who participated in the present study was high school and junior high school principals or assistant principals. Examination of Table 1 reveals that nearly half (44%) of the participants were principals; the others (56%) were assistant principals. Most worked in a senior high (56%) with between 1,000 and 2,500 students (50%), in a district having a total enrollment over 10,000 (48%). A majority of the administrators reported the Master of Science degree as their highest completed educational level (64%) and were from Iowa (90%). Information about two other variables describing the sample is presented in Table 2: years of administrative experience (\bar{X} =12 years), and the influence that the administrator's opinion has on personnel decisions at his/her school (\bar{X} =64% influence).

Table 1
 Sample Profile
 (Categorical Data)

	Number	Percent
Position		
Principal	35	43.8
Assistant principal	45	56.2
Total	80	100.0
Level		
High school	45	56.2
Junior high school	35	43.8
Total	80	100.0
District enrollment		
250- 1,000	6	7.5
1,000- 5,000	10	12.5
5,000-10,000	26	32.5
10,000-50,000	38	47.5
Total	80	100.0
Building enrollment		
70- 150	0	0.0
150- 500	15	18.8
500-1,000	25	31.3
1,000-2,500	40	49.9
Total	80	100.0
Degree		
M.S.	51	63.7
M.S. +45 or specialist	23	28.8
Ed.D. or Ph.D.	6	7.5
Total	80	100.0
State		
Iowa	72	90.0
Georgia	1	1.2
Illinois	3	3.8
New York	4	5.0
Total	80	100.0

Table 2
 Sample Profile
 (Continuous Data)

	Range	Mean	Standard deviation	Number
Years of administrative experience	1- 29	11.68	6.59	80
Influence on personnel decisions (percent)	0-100	63.83	21.01	80

Reliability Estimates

In order for the LISREL IV procedure to adjust for measurement error, it was necessary to have available error variances or reliabilities for the variables, or to estimate error variances or reliabilities of the data from the study conducted. The researcher adopted a procedure used by Adb-Ella et al. (1981) to obtain reliabilities for the 15 variables in this study. For variables which were measured as a single item -- years of administrative experience, building enrollment, district enrollment, position, and influence -- available reliabilities were used to provide estimates of measurement error. Kelley (1973), Otto and Featherman (1975), and Fuller and Hidioglou (1978) provided available reliabilities for these variables. For seven of the eight variables measured by more than one item -- knowledge, open mindedness, evaluation experience, evaluator quality, usefulness of the information, usefulness of the process, and overall impact -- coefficient alpha was used as an internal consistency measure of reliability. The other

scale, uses, was not hypothesized to be internally consistent. Therefore, a panel of judges estimated the test-retest reliability of this scale. The treatment variables were assumed to be without measurement error.

Information in Table 3 summarizes descriptive and reliability results.

Because of adjustments made by the researcher, the statistics for the knowledge scale should be discussed at this time. The original knowledge quiz consisted of 10 items. The sample results produced one item which correlated negatively with the total (item 3, $r=-.33$), and three items with unacceptable item discrimination (item 5, $p=0.00$; item 4, $p=.81$; item 8, $p=.02$). Therefore, the researcher eliminated these four items from the final data analysis, producing the statistics shown in Table 3.

Least Squares Regression Analysis

The researcher investigated five hypotheses relating various sets of predictor variables to the five criterion variables. Ordinary multiple regression results corresponding to each hypothesis are presented. Later in this chapter, the results of the test of the overall model, corrected for measurement error, are described.

Hypothesis 1. There is no linear relationship between the the resume treatment, years of administrative experience, evaluation experience, decision making responsibility, and the information treatment as predictors of the criterion, quality of the evaluator.

Results relating to Hypothesis 1 are presented in Table 4.

Table 3
Summary Statistics for Measures

Theoretical concept	Empirical measure	Mean	Variance	Reliability
Evaluator background	Resume treatment (RTRT) ^a	1.5	.25	1.00
Open mindedness	Rokeach dogmatism scale (ROKEACH)	131.8	461.61	.79
Knowledge of program evaluation	Total on knowledge quiz (KNOW)	1.26	1.44	.57
Administrative experience	Years as a school administrator (YRS)	11.69	43.53	.93 ^b
Evaluation experience	Total on evaluation activities scale (EVALEXP)	9.09	24.56	.82
Decision making responsibility (DMR)	Composite of position (POS), Building enrollment (BENR), District enrollment (DENR), and Influence on personnel decisions (INF)	13.59	4.08	.72 ^b
Decision maker awareness	Information treatment (ITRT)	.50	.25	1.00
Evaluator quality	Total evaluator quality scale (EQ)	47.05	53.95	.87
Usefulness of information	Total on information scale (UI)	14.19	7.60	.75
Usefulness of process	Total on process scale (UP)	13.89	7.24	.73
Number of uses	Total uses scale (USES)	5.23	2.96	.70 ^c
Overall impact	Total usefulness scale (OAIMPACT)	10.6	4.72	.90

^aLabels in parentheses used in future tables, figures, and discussion.

^bAvailable reliability estimates for a similar scale (Abd-Ella *et al.*, 1981).

^cEstimated by a panel of judges.

Table 4
The Relationship between Predictor
Variables and Evaluator Quality

	Zero-order correlations	Regression coefficients	
		Unstandardized	Standardized
Resume treatment	.162	1.656	.113
Administrative experience	-.192*	-0.205	-.185
Evaluation experience	-.091	.042	.028
Decision making responsibility	-.170	-.604	-.166
Information treatment	.090	.845	.058
Constant		54.373	
R ²		.085	

*Significant at .05 level.

The correlation coefficients indicate that the variable, years of administrative experience, is correlated with ratings of evaluator quality and that this relationship is negative. This relationship indicates that the more experience the administrator had the lower the administrator rated the quality of the evaluator. The multiple regression coefficients indicate that the give variables explain 8.5 percent of the variance in evaluator quality. None makes a significant contribution to the regression equation. The researcher failed to reject Hypothesis 1.

Hypothesis 2. There is no linear relationship between years of administrative experience, evaluation experience, knowledge, and the information treatment as predictors of the criterion, uses.

Results relating to Hypothesis 2 are presented in Table 5.

Table 5
The Relationship between
Predictor Variables and Uses

	Zero-order correlations	Regression coefficients	
		Unstandardized	Standardized
Administrative experience	.107	.025	.094
Evaluation experience	.167	.040	.114
Information treatment	.00	.204	.060
Knowledge	.266**	.369*	.257
Constant		4.010	
R ²		.097	

*Significant at .05 level.

**Significant at .01 level.

Multiple regression results indicate that the four variables explain 9.7 percent of the variance in uses. The others fail to make significant contributions. Therefore, the researcher failed to reject Hypothesis 2. However, the correlation coefficients indicate that knowledge is positively correlated with uses. Knowledge makes a significant contribution to the regression equation. This implies that the likelihood of using evaluation information increases as the decision maker's knowledge about program evaluation increases.

Hypothesis 3. There is no linear relationship between evaluation experience, decision making responsibility, the information treatment, knowledge, and open mindedness as predictors of the criterion, usefulness of the evaluation process.

Findings related to Hypothesis 3 are presented in Table 6.

Table 6
The Relationship between Predictor
Variables and Usefulness of Process

	Zero-order correlations	Regression coefficients	
		Unstandardized	Standardized
Evaluation experience	.007	.023	.042
Decision making responsibility	-.186*	-.202	-.152
Information treatment	.033	.216	.040
Knowledge	.111	.147	.065
Open mindedness	.195*	.021	.170
Constant		18.935	
R ²		.070	

*Significant at .05 level.

The correlation coefficients indicate that open mindedness and decision making responsibility are correlated with usefulness of process. The relationship between decision making responsibility and usefulness of process is negative, indicating that the more responsibility a decision maker has the lower the decision maker rated the evaluation process. The positive relationship between the Rokeach score and usefulness of process indicates that open minded decision makers rated the process higher than closed minded decision makers. The multiple regression results indicate that the five variables explain 7.0 percent of the variance in usefulness of process. Also, none of the five makes

a significant contribution to the total. Therefore, the researcher failed to reject Hypothesis 3.

Hypothesis 4. There is no linear relationship between evaluation experience, decision making responsibility, knowledge, open mindedness, and the information treatment as predictors of the criterion, usefulness of evaluation information.

Findings related to Hypothesis 4 are presented in Table 7.

Table 7
The Relationship between Predictor Variables
and Usefulness of Information

	Zero-order correlations	Regression coefficients	
		Unstandardized	Standardized
Evaluation experience	.031	.043	.076
Decision making responsibility	-.186*	-.187	-.137
Information treatment	.059	.470	.086
Knowledge	.146	.209	.091
Open mindedness	.264**	.031*	.242
Constant		19.930	
R ²		.109	

*Significant at .05 level.

**Significant at .01 level.

The correlation coefficients indicate that two variables, decision making responsibility and open mindedness, are correlated with usefulness of evaluation information. A high score on the Rokeach Dogmatism Scale, which indicates open mindedness, corresponds with a high rating

of usefulness of information. Decision making responsibility is negatively related to usefulness of information. This negative correlation suggests that the more responsibility a decision maker had the lower the decision maker rated the utility of the information in the evaluation report. Multiple regression results indicate that 10.9 percent of the variance in usefulness of information can be accounted for by these five variables. Open mindedness contributes significantly to the regression equation. The others fail to make a significant contribution. The researcher failed to reject Hypothesis 4.

Hypothesis 5. There is no linear relationship between evaluator quality, uses, usefulness of process, and usefulness of information as predictors of the criterion, overall impact.

Findings related to Hypothesis 5 follow in Table 8.

The correlation coefficients indicate that all four predictors are positively related to the overall impact of the evaluation. That is, high ratings of evaluator quality, number of uses, usefulness of process or information corresponded with high ratings of overall impact. Multiple regression results indicate that the equation containing the four predictors explains 71 percent of the variance. Therefore, the null hypothesis is rejected and the researcher concludes that there is a predictable linear relationship between the criteria and overall impact. Three of the four predictors contribute significantly to overall impact; uses does not.

The correlation matrix in Table 9 is a display of the relationships among the seven predictor and five criterion variables. Relationships that were not evidenced in the presentation of the regression results are now shown.

Table 8
 The Relationship between Predictor
 Variables and Overall Impact

	Zero-order correlations	Regression coefficients	
		Unstandardized	Standardized
Uses	.373***	-.166	-.132
Evaluator quality	.600***	.052*	.176
Usefulness of process	.793***	.327**	.405
Usefulness of information	.794***	.331**	.419
Constant		-.213	
R ²		.710**	

*Significant at .05 level.

**Significant at .01 level.

***Significant at .001 level.

Although their relationships were not hypothesized, analysis of the data reveals a positive relationship between open mindedness (ROKEACH) and evaluator quality (EQ) ($r=.33$, $p \leq .001$). Decision making responsibility (DMR) and the overall impact (OAIMPACT) of the evaluation are negatively related ($r=-.20$, $p \leq .05$). Also, the resume treatment (RTRT) is significantly related to the usefulness of the process (UP) ($r=.24$, $p \leq .05$). This relationship was investigated further. A multiple classification analysis of variance measuring the effect of the two treatments on ratings of the usefulness of the process produced the results as found in Table 10. Examination of the

Table 9
Correlation Coefficients

	EQ	UI	UP	USES	OAIMPACT	RTRT	ROKEACH
EQ							
UI	.61***						
UP	.58***	.83***					
USES	.52***	.52***	.48***				
OAIMPACT	.60***	.79***	.79***	.37***			
RTRT	.16	.15	.24*	.10	.12		
ROKEACH	.33***	.26**	.19*	.17	.10	.07	
KNOW	.02	.15	.11	.27**	.08	-.05	.11
YRS	-.19*	-.02	.001	.11	.01	-.18	.21*
EVALEXP	-.09	.03	.01	.17	-.003	-.13	.04
DMR	-.17	-.19*	-.19*	-.01	-.20*	-.11	.15
ITRT	.09	.06	.03	.00	.09	.00	.06
BENR	-.02	-.11	-.15	.05	-.13	-.13	-.04
DENR	-.26**	-.26**	-.18	-.06	-.18*	-.16	-.16
POS	.04	-.12	-.09	-.09	-.01	.07	-.04
INF	-.11	.10	.04	.07	-.07	-.01	-.06

*Significant at .05 level.

**Significant at .01 level.

***Significant at .001 level.

KNOW	YRS	EVALEXP	DMR	ITRT	BENR	DENR	POS	INF
-.07								
.16	.32**							
.15	-.06	.16						
-.05	-.10	-.32**	-.13					
-.14	-.07	.17	.80**	-.13				
-.21*	.09	.01	.63***	-.08	.34***			
-.27**	-.45***	-.32**	.60***	.07	.49***	.27**		
.30**	.29**	.47***	.01	-.13	-.19*	-.26**	-.52***	

cell means indicates that the group that read the evaluator's resume ($\bar{X}=14.52$) rated the usefulness of the process higher than the group that read the researcher's resume ($\bar{X}=13.25$).

LISREL IV Analysis

Although examination of the correlation matrix in Table 9 indicates that the hypothesized model would not yield accurate predictions of overall impact, the original model was tested using the LISREL technique. This was done for two reasons. First, such an investigation enabled the researcher to compare the path analysis and least squares results. Also, it provided empirical information to guide the revision

Table 10

Analysis of Variance: Usefulness of Process
by Resume Treatment and Information Treatment

Source	df	SS	MS	F
ITRT	1	.612	.612	.087
RTRT	1	32.512	32.512	4.616*
Interaction	1	3.613	3.613	.513
Residual	76	535.25	7.043	
Total	79	571.982	7.24	

*Significant at .05 level.

of the model. The original model displayed in Figure 4 was tested. The labels used in Figure 4 were defined in Table 3.

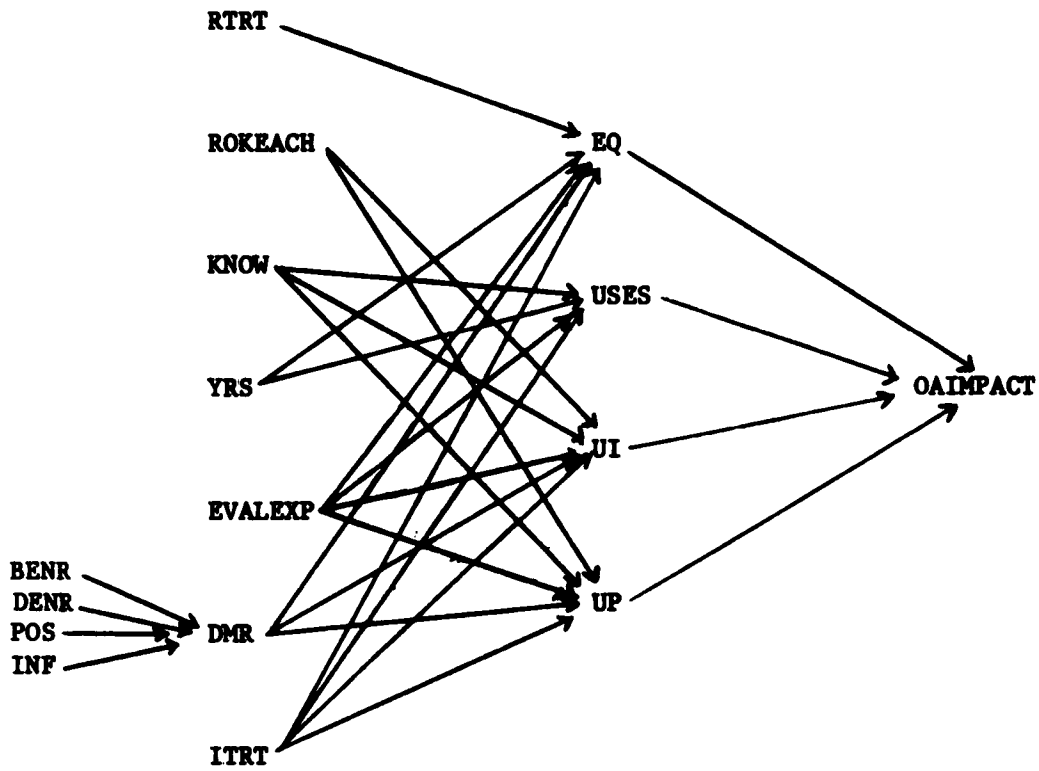


Figure 4. Original evaluation utilization measurement model

Before the measurement model could be analyzed, one change had to be made. The high correlation between UP and UI ($r=.83$, $p \leq .001$) along with their high reliabilities, UP ($\alpha=.73$), UI ($\alpha=.75$) caused the measurement of these variables to be overcorrected, and made solution of the equations impossible. Therefore, the researcher determined that the two scales should be combined into one composite (USEFUL) by the LISREL program, and the solution sought.

The solution obtained was only a partial one, indicating that the relationships among the variables greatly conflicted with the hypothesized relationships in the model. Based on the test of goodness of fit for this partial solution, the researcher concluded that the data did not fit the original model ($\chi^2_{63}=183.22$, $p \leq .0001$). With assistance from Dr. Richard Warren, Director of the Research Institute for Studies in Education at Iowa State University, the researcher studied the LISREL estimates for the path coefficients, their corresponding t-values, the residual matrix (see Appendix C for these), and the original correlation matrix to revise the original model. In this discussion the term, path coefficients, refers to raw regression coefficients adjusted for measurement error.

It should be mentioned here that theory and past research findings were studied to develop the original model. Examination of the data of the present study guided the model revisions. However, if an illogical relationship was suggested by the data, the researcher rejected its inclusion into the model, thus maintaining the model's integrity.

Results of the test of the original model indicated that the composite, DMR, was not behaving predictably. Examination of the correlation matrix indicated that two of the items in the composite, BENR and DENR, were not consistently associated with other variables in the matrix. That is, the items in a composite should be more highly intercorrelated than they are correlated with other variables, and that the correlations among the variables in the composite with other variables should be similar. In this data set, POS and INF were related to other decision maker characteristics; DENR and BENR were not. Therefore, the composite DMR was revised to include only POS and INF.

A second change made in the model was to include ROKEACH as a predictor of EQ. The residual covariance between these two variables was 42.29, whereas the correlation was .33 ($p \leq .001$).

The third change in the model was major. There were high correlations among the criterion variables. Also, the LISREL results indicated that the first level criteria (USES, EQ, and USEFUL) are accurate predictors of OAIMPACT. That is, their path coefficients were significantly different than zero. This conflicted somewhat with the multiple regression results presented in Table 8. The variable, USES, was behaving differently in the two analyses. In addition, the residual covariances between EQ, USES, and USEFUL, as found in the LISREL results, indicated that the model was not reflecting the relationships in the data. The correlation between USES and OAIMPACT was smaller than the other correlations in this

group. Therefore, it was hypothesized that USES should be a second level criterion caused by EQ and USEFUL. This suggests that the relationship between USES and OAIMPACT is due to the fact that they both were so strongly associated with EQ and USEFUL.

The revised measurement model reflecting these three changes is shown in Figure 5.

The goodness of fit test for this model indicates that the data fit this revised model better than they fit the original model ($\chi^2_{36} = 72.54, p \leq .0003$). However, these results suggested another revision. The estimated path coefficients supported the hypothesis that EQ causes USES, and that USEFUL causes OAIMPACT (see Appendix D). There was no support for the causal relationship between USEFUL and USES, or EQ and OAIMPACT. Therefore, the unsupported relationships were dropped for the second revision.

KNOW and YRS were also determined to cause USES. KNOW was shown to cause USEFUL. Although the path coefficient for YRS when predicting EQ was not statistically significant, it was more closely related than the other hypothesized predictors for EQ and therefore was retained in the second revision.

Also, the problems with the original model's residual matrix discussed earlier were not evidenced in this first revised model. The second revision is displayed in Figure 6.

Clearly, this model is much reduced from the original. The results of testing this model indicate that it fits the data less accurately

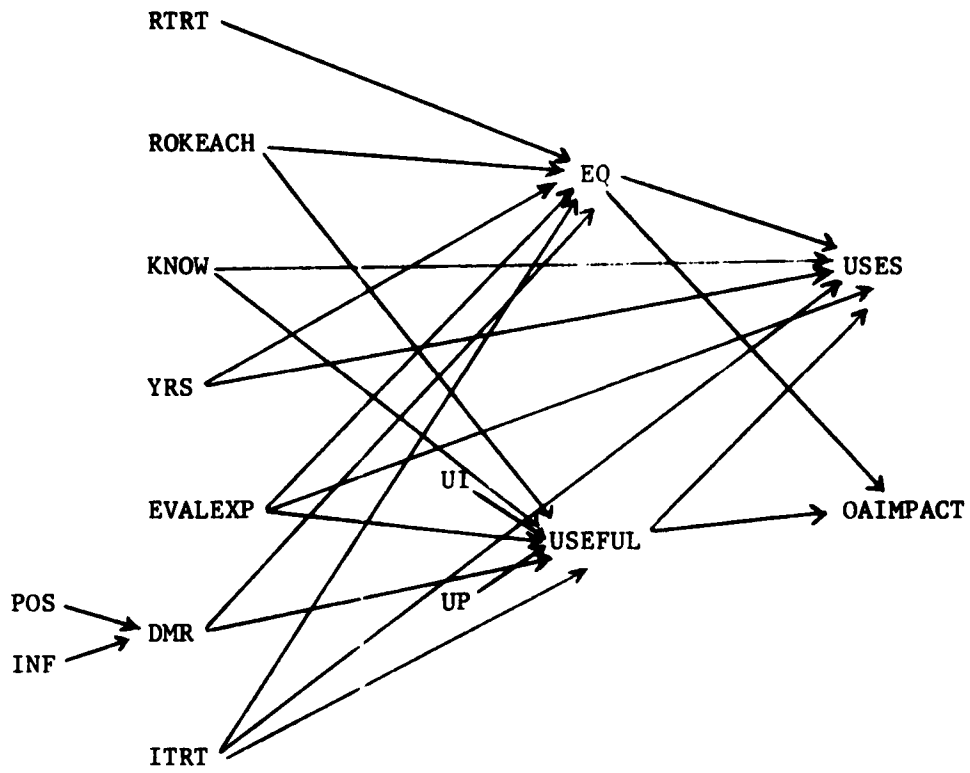


Figure 5. Revised measurement model (1)

than the first revision did ($\chi^2_{49}=109.44$, $p=0.00$). Therefore, with an additional 13 degrees of freedom, the chi-square value increased by 37.

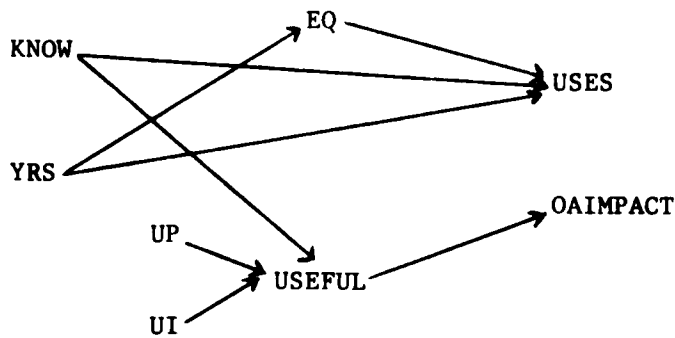


Figure 6. Revised measurement model (2)

The problems caused by ROKEACH in the residual matrix of the original matrix surfaced again (see Appendix E). Also, the residuals between EQ and USEFUL increased problematically. These facts suggested a further revision replacing the ROKEACH variable in the model and replacing the arrows between USEFUL and USES and between EQ and OAIMPACT. Also, the relationship between RTRT and USEFUL had never been investigated in any analysis, so it was included in the third and final revision which is presented below in Figure 7.

The goodness of fit test for this final model indicates that this revision improved the fit of the data over revision 2 ($\chi^2_{44}=90.22$, $p\leq 0.0001$). However, it was not an improvement from revision 1. Seven of the 11 path coefficients estimated were statistically significant (see

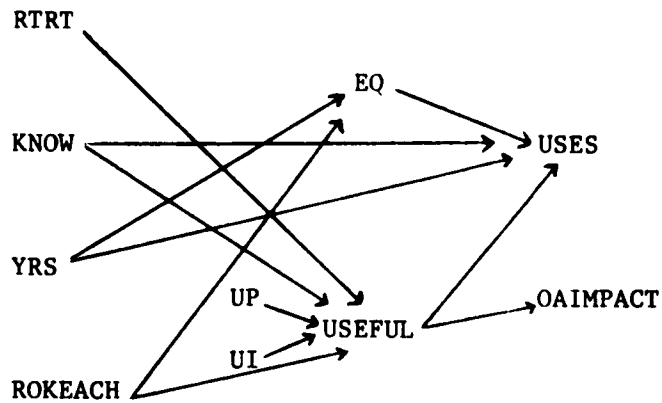


Figure 7. Revised measurement model (3)

Appendix F). The residual matrix was comparable to that resulting from revision 1.

Therefore, the researcher concluded that the data from this sample did not fit any of the four models well. However, there was support for the theoretical model suggesting that evaluator, decision maker, program and context, and report and process characteristics affect the impact of an evaluation.

CHAPTER V. SUMMARY AND CONCLUSIONS

Summary

The purpose of the present study was to investigate the effect that evaluator and decision maker characteristics had on the impact of a program evaluation. The need for research in this area was discussed in the introductory chapter. Many professionals involved in evaluation activities were dissatisfied with the minimal effect evaluations had on program planning and decisions. Therefore, researchers began to investigate components of useful evaluations.

The present research was founded on a model including four components thought to affect evaluation utilization. Only two of the four, evaluator and decision maker, were investigated in this study. The four factors in the model, as well as the seven predictor variables studied in this investigation, are highlighted in the literature review section.

The method used by the researcher to investigate this problem was a simulated educational program evaluation. The simulation materials described in the methodology chapter are included in the appendices. Eighty secondary school administrators were the decision makers participating in this study who read the simulated materials and responded to questions measuring the usefulness criteria. In addition, the administrators supplied background information, as five of the variables examined were decision maker characteristics.

The relationships between the seven predictor and five criterion variables, as shown in Figure 4, were tested using both least squares

regression and path analysis techniques. In addition, the overall model was tested by means of the path analysis procedure, LISREL IV. These findings, as well as the results of three model revisions, were reported in the findings chapter.

The findings section is summarized next. Correlation and ordinary multiple regression results support the following relationships. The amount of knowledge a school administrator had about program evaluation predicted the number of evaluation uses selected by the administrator. That is, administrators with more knowledge chose more uses for the simulated evaluation. In addition, scores on the Rokeach Dogmatism Scale predicted ratings of usefulness of the evaluation information and correlated with ratings of usefulness of the evaluation process and of evaluator quality. Therefore, open minded decision makers perceived the process and information as more useful than did closed minded decision makers. Also, open minded principals rated the quality of the evaluator higher than closed minded principals did.

Two other decision maker characteristics were correlated with the usefulness criteria. Years of administrative experience was negatively correlated with evaluator quality suggesting that more experienced principals graded the evaluator lower than less experienced principals did. Also, there was an inverse relationship between decision making responsibility and the criteria, usefulness of information, usefulness of process, and overall impact. Therefore, the more responsibility an administrator had the lower he/she rated the process, information, and impact of the evaluation. Perhaps administrators

with many responsibilities felt that they could make good decisions without an evaluator's services.

The background of the evaluator was correlated with one criterion, usefulness of process. Evidence was presented to indicate that the process used by evaluator was rated higher than the process used by the researcher. Decision makers may have identified more with the evaluator than with the researcher.

There were no statistically significant relationships between evaluation experience or the information treatment and any of the dependent variables. Therefore, these two could not predict any of the five criteria measuring usefulness.

Multiple regression results supported the hypothesized relationship between the overall impact of the evaluation and the four first level criteria, evaluator quality, number of uses, usefulness of information, and usefulness of the process. This result indicates that inquiring about the evaluator, uses, usefulness of the process and information is an accurate method for determining the overall impact of an evaluation.

The LISREL results supported fewer relationships than the multiple regression results did. The predictive relationship between the knowledge variable and the number of uses selected by the decision maker was again evidenced. Also, the open mindedness variable proved important. Rokeach scores predicted both evaluator quality and the useful composite. Years of experience was the only other decision maker characteristic predicting a criterion. Years predicted number

of uses. That is, administrators with more experience selected more uses than administrators with less experience selected.

Although the correlational analysis suggested the importance of decision making responsibility and evaluator background, neither of these had statistically significant path coefficients in any of the LISREL analyses. It must be noted that the decision maker variable used in the LISREL analysis consisted of two variables: position and influence. Also, the LISREL program weighted each variable forming the decision making responsibility composite based on the covariances among the variables. For the correlation analysis four variables formed the decision making responsibility composite: position, influence, building enrollment, and district enrollment. They were standardized and summed without any weighting. If building and district enrollments had been continuous variables, the researcher would have more confidence in the correlation results. Also, the advantages of the LISREL procedure suggest to the researcher the efficacy of the path analysis results that decision making responsibility had no effect on perceptions of evaluation usefulness.

The other independent variable producing conflicting results was evaluator background. Whereas the correlation between background and evaluator quality suggested that the process used by an evaluator was more useful than that of a researcher, when usefulness of process and usefulness of information were combined to form the useful composite by the LISREL procedure, there was no evidence of such a relationship. One reason for this might be that when the composite

was formed the effect of the process variable was masked by the information variable. Therefore, the researcher has confidence in the correlation results that evaluator background does affect usefulness of process ratings. In terms of predicting overall impact, the multiple regression results are clear and simple. A regression equation containing uses, evaluator quality, usefulness of information, and usefulness of process accurately predicted overall impact. The LISREL results for this same relationship are less succinct. First, the effect of the useful composite must be considered. When the final measurement model was tested, the useful composite predicted both number of uses and overall impact. However, evaluator quality predicted the number of uses only; it made no contribution to predicting overall impact.

Subsequent discussion of the findings will be based on the results of the third revision of the measurement model. The major reason for this decision was that the nature of the number of uses criterion in the LISREL analysis is logical and is supported by the data. That is, due to the evidence suggesting that the correlation between uses and overall impact is due to their high correlations with evaluator quality, usefulness of information, and usefulness of process, number of uses was moved to the same level as overall impact.

Lastly, none of the four models evaluated using the goodness of fit test accurately fit the data. However, results outlined above provide support for the importance of certain evaluator and decision maker variables in determining the usefulness of an evaluation.

Therefore, although no measurement model could be adopted in its entirety, evidence supports the evaluation utilization model.

Discussion

The criteria employed in the present study reflected a broad definition of evaluation utilization (Patton, 1978; Alkin et al., 1979). Therefore, the impact referred to in this discussion includes long-range, indirect effects of evaluation findings as well as immediate, direct uses. This is most clearly evidenced in the list of possible uses from which administrators selected. The complexity of measuring evaluation usefulness was also accounted for in the present study by having five different measures of impact rather than only one. Dickey (1980) is an example of another study which attempted to capture the multivariate nature of utilization. Therefore, the researcher assumed that decision makers use input other than evaluation results when making decisions. She also assumed that major, immediate decisions are the exception rather than the rule. Thus, using these five measures of usefulness rather than only one, resulted in findings comparable to what one would find in an actual field study. Hence, the findings and results of this study are valuable. It also should be noted that for the path analysis, two criteria, usefulness of process and usefulness of information, were combined to form the useful composite.

The relationship between evaluator background and ratings of evaluator quality add to the findings of other studies measuring the effects of the evaluator component of the model used in the present

study. Archibald (1970) and Braskamp et al. (1978) found that the reputation or title of the person conducting the evaluation was important. Whereas Braskamp et al. found ratings of the objectivity of a researcher highest, in the present study, the process used by the evaluator with an evaluation background was rated higher. Background made no difference in ratings of evaluator quality in the present study. Further investigation determined that the two evaluator background treatment groups rated the objectivity of the evaluator the same ($t=.37$, $p=.71$). Therefore, evidence provided by the current study supports the fact that evaluator characteristics influence evaluation usefulness, but it is unclear if background or title are the most important characteristics. Hence, there is support for the inclusion of the evaluator factor in a theoretical evaluation utilization model. However, no definitive statement can be made recommending one background over another.

The first variable discussed in the decision maker category, open mindedness, predicted both evaluator quality and the usefulness composite. That is, open minded principals rated the evaluator, information and process variables, higher than closed minded principals did. This supports the idea that open minded decision makers respond positively to evaluation activities and results. This implies that open minded administrators might be more cooperative than closed minded administrators. Practically, recognizing a dogmatic decision maker is the problem facing evaluators. As described in Chapter II, dogmatic decision makers make rapid decisions based on relatively

little information. For a clue as to whether or not a decision maker is open or closed minded, evaluators could probe decision makers about processes used to make previous decisions. The next question which needs study is, "What are the most effective evaluation strategies to use when working with closed minded decision makers?" This finding compares with the importance of decision making orientation (Stecher, 1981) and management style (Davis, 1981). The decision making responsibility variable investigated in the current study did not affect any dependent variable. However, the importance of open mindedness, decision making orientation, and management style combined with the problems associated with the decision making responsibility variable in the present study, leads to the conclusion that the decision maker factor in the evaluation utilization model is imperative. These findings highlight a statement reported in Chapter II:

...the evaluator must realize that the value system, philosophical orientation, political ambitions, and personality makeup, and 'who knows what else' affect the extent to which evaluation findings are utilized (Alkin, 1975, p. 201).

Therefore, it is strongly recommended that evaluators learn as much as they can about the decision makers with whom they work, and consider the decision maker's attitudes and values in all stages of planning, implementing, and disseminating the evaluation. Evaluators can do this by involving decision makers in the evaluation process, seeking input from them, and keeping them informed of findings during all phases of the evaluation and by writing results in a readable style.

The open mindedness of a decision maker is something that an evaluator cannot change. However, the next decision maker variable, decision maker knowledge of program evaluation, is a variable over which the evaluator has some control. Results of the current study indicate that the more knowledge a decision maker has, the more uses he/she can envision for the evaluation findings. This adds to the findings of Brown et al. (1980) who found that knowledge made a difference in one study, but made no difference in another study. The knowledge variable used by Brown et al. (1980) was a treatment variable comparable to the information treatment investigated in the present study. The treatment variable had no effect on any of the usefulness criteria. That is, reading an essay concerning the similarities and differences between research and evaluation did not influence perceptions of usefulness. One reason for this might be that this specific content does not affect the number of uses decision makers selected. Considering the low reliability of the knowledge quiz developed for this study and the measurement problems associated with it, the researcher can conclude that knowledge may affect ratings of usefulness. However, this question clearly needs additional study. This variable especially deserves attention because it is an "alterable" variable. That is, the evaluator could take time to educate decision makers about evaluation, if there was evidence that increased knowledge resulted in increased utilization. Also, administrator training programs could include program evaluation in the curriculum if knowledge proved to be an important variable.

Years of administrative experience was included in the current study to replicate Lorenzen and Braskamp (1978). Whereas they found experience to be unrelated to usefulness, the current findings indicate that years of experience predicts the number of uses of an evaluation selected by administrators. The researcher hypothesized that the evaluation experience of a decision maker would be more sensitive to ratings of usefulness than years of experience. Current results offer no support for this hypothesis. Like knowledge, years of experience warrants further investigation. Years of experience, however, is not an "alterable" variable; therefore, for practical significance, further investigation of it is not as important as additional research on knowledge.

In addition to the relationships between predictors and criteria discussed above, the model predicted relationships among the dependent variables. The correlation coefficients among the five usefulness variables were all statistically significant and positive (see Table 9). However, the regression and path coefficients present confusing results. It appeared that the relationship between number of uses and overall impact was due to the common factor they shared with the evaluator quality variable and the usefulness composite. Therefore, during the model revision stage of the analysis, number of uses was hypothesized to be caused by evaluator quality and usefulness. This relationship was supported when subsequent revisions were tested. However, the predictive ability of quality and usefulness did not remain dependable when overall impact was the dependent variable. The significance and

predictability of number of uses suggest one method for evaluators to determine the impact of an evaluation. The evaluator could follow up post evaluations at several time periods, for example, soon after the results were presented, six months later, and one year after the completion of the evaluation. Inquiries about the uses to which the evaluation information was put could be made at these times. The checklist developed by Dickey (1981) could be employed. Additional uses could also be solicited. This suggestion could also be tested empirically to determine if those evaluations used for many purposes were also rated as having the greatest impact. However, this requires further investigation before any strong recommendation can be made.

Implications for evaluators

The results of the present study, combined with prior findings, indicate the importance of the personality of the decision maker in determining the usefulness of the evaluation. The researcher, therefore, concurs with the purpose and involvement recommendation discussed in Chapter II, as well as the personal factor (Patton, 1978) and the orientation and administrator style components in Alkin *et al.*'s (1979) framework. That is, the evaluator must take time to know the decision makers, and plan the evaluation accordingly, as well as to include the decision maker in all phases of the evaluation process.

In addition, although the findings related to the knowledge variable were not conclusive, it probably would do no harm for an evaluator to unobtrusively assess an administrator's knowledge of program evaluation and then attempt to educate the decision maker on

important concepts in the area of program evaluation. The relationship between knowledge and the usefulness variables was positive, offering some evidence supporting this recommendation.

Implications for decision makers

Assuming that quality evaluation information reduces a decision maker's field of uncertainty and therefore improves decision making efforts, program administrators could increase their effectiveness by maximizing the use to which they put evaluation results. They can do this by being honest with the evaluator about their purposes and motivations, by cooperating in planning, implementing, and disseminating efforts, by being open to new or surprising information about the program, and by becoming informed about the field of program evaluation.

Suggestions for further research

The evaluation factor of the evaluation utilization model needs additional investigation to determine which evaluator characteristics positively affect utilization. This information would be valuable for those involved with training evaluators. Perhaps personality and knowledge variables of the evaluator, similar to those of the decision maker, would prove useful.

The two decision maker variables needing further study are knowledge of program evaluation and experience. Before any strong recommendations can be made concerning the importance of decision maker knowledge of program evaluation, clearer results must emerge.

However, it is nearly certain that decision maker characteristics affect the impact of evaluative information.

In addition to these suggestions, the model developed, tested, and revised, can assist future researchers refine an evaluation utilization model.

The purpose of this research was to test theoretical relationships between evaluator and decision maker characteristics and the impact of an evaluation. Evidence was provided in support of the theoretical evaluation utilization model developed. In addition, the researcher provided suggestions for further research, implications for evaluators, and implications for decision makers.

REFERENCES

- Abd-Ella, M. M., Hoiberg, E. O., & Warren, R. D. Adoption behavior in family farm systems: An Iowa study. Rural Sociology, 1981, 46, 42-61.
- Alkin, M. C. Evaluation: Who needs it? Who cares? Studies in Educational Evaluation, 1975, 1, 201-212.
- Alkin, M. C. Evaluation use in local schools. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Alkin, M. C., Daillak, R., & White, R. Using evaluations. Beverly Hills: Sage Publications, 1979.
- Anderson, S., & Ball, S. The profession and practice of program evaluations. San Francisco: Jossey-Bass Publishers, 1978.
- Apling, R. How evaluators can help program managers improve their use of information from tests and evaluations. Paper presented at the American Educational Research Association annual meeting, Los Angeles, 1981.
- Archibald, K. Alternative orientations to social science utilization. Social Science Information, 1970, 9(2), 7-34.
- Becker, H. The effect of presentation style on audience reaction to a program evaluation report. Evaluation News, 1981, 2(1), 77-78.
- Blalock, H. M., Jr. Some implications of random measurement error for causal inferences. American Journal of Sociology, 1965, 71(July), 37-47.
- Bohrnstedt, G. W., & Carter, T. M. Robustness in regression analysis. In H. L. Costner (Ed.), Sociological methodology. San Francisco: Jossey-Bass Publishers, 1971.
- Boruch, R., Cordray, D., Pion, G., & Leviton, L. A mandated appraisal of evaluation practices: Digest of recommendations to the Congress and the Department of Education. Educational Researcher, 1981, 10(4), 10-13.
- Braskamp, L. A., & Brown, R. D. (Eds.). Utilization of evaluative information. New directions for program evaluation (Number 5). San Francisco: Jossey-Bass Publishers, 1980.

- Braskamp, L. A., Brown, R. D., & Newman, D. L. The credibility of a local educational program evaluation report: Author source and client audience characteristics. American Educational Research Journal, 1978, 15, 441-450.
- Brown, R. D., Braskamp, L. A., & Newman, D. L. Evaluator credibility as a function of report style. Evaluation Quarterly, 1978, 2, 331-340.
- Brown, R. D., Newman, D. L., & Rivers, L. Perceived need for evaluation and data usage as influences on an evaluation's impact. Educational Evaluation and Policy Analysis, 1980, 2(5), 67-73.
- Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56(March), 81-105.
- Ciarlo, J. A. (Ed.). Utilizing evaluation. Sage research progress series in evaluation (Vol. 6). Beverly Hills: Sage Publications, 1981.
- Conner, R. F. Measuring evaluation utilization: A critique of different techniques. In J. A. Ciarlo (Ed.), Sage research progress series in evaluation (Vol. 6). Beverly Hills: Sage Publications, 1981.
- Cronbach, L. J. Course improvement through evaluation. Teachers College Record, 1963, 68, 672-683.
- Davis, D. The effects of administrative style on evaluation utilization. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Davis, F. R., & Salasin, S. D. The utilization of evaluation. In E. L. Struening & M. Guttentag (Eds.), Handbook of evaluation research (Vol. 1). Beverly Hills: Sage Publications, 1975.
- Deats, T. Moving and using information. Teachers College Record, 1974, 75, 383-393.
- Dickey, B. Utilization of evaluations of small scale innovative educational projects. Educational Evaluation and Policy Analysis, 1980, 2(6), 65-77.
- Evers, F. T. A systematic approach to organizational analysis: The case of farmer cooperatives (Doctoral dissertation, Iowa State University, 1979). Dissertation Abstracts International, 1980, 40, 4270A. (University Microfilms No. 8000131.)

- Ferman, L. A. Some perspective on evaluating social welfare programs. The Annals of the American Academy of Political and Social Science, 1969, 385(September), 143-156.
- Fuller, W. A., & Hidioglou, M. A. Regression estimates after correcting for attenuation. Journal of the American Statistical Association, 1978, 73(March), 99-104.
- Guba, E. G. The failure of educational evaluation. Educational Technology, 1969, 9(5), 29-38.
- Haenn, J., & Owens, T. Utilization frameworks for evaluation reporting. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Horowitz, J. Utilization of various evaluation activities in urban schools with specially funded programs. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Joint Committee on Standards for Educational Evaluation. Standards for evaluations of educational programs, projects, and materials. New York: McGraw-Hill Book Company, 1981.
- Joreskog, K. G., and Sorbom, D. LISREL IV users guide. Chicago: National Educational Resources, Inc., 1978.
- Kelley, J. Causal chain models for the socioeconomic career. American Sociological Review, 1973, 38, 481-493.
- Kennedy, M. How evaluations can help evaluators. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Locatis, C., Smith, J., & Blake, V. Effects of evaluation information on decisions. Evaluation Review, 1980, 4, 809-821.
- Lorenzen, G., & Braskamp, L. Comparative influence of political, cost/benefit, and statistical information on administrative decision making. Evaluation and Program Planning, 1978, 1, 235-238.
- MacCrimmon, K. R., & Taylor, R. Decision making and problem solving. In M. Dunnett (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally Publishing Company, 1976.
- Mann, J. The outcome of evaluative research. In C. H. Weiss (Ed.), Evaluating action research: Readings in social action and education. Boston: Allyn and Bacon, Inc., 1972.

- National Advisory Council on the Education of Disadvantaged Children. 1979 Annual Report to the President and the Congress. Washington, D.C.: Author, 1979.
- Neumann, W. Evaluators and principals: The aim is self-reliance. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. SPSS: Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, Inc., 1971.
- Otto, L. B., & Featherman, D. L. Social structural and psychological antecedents of self-estrangement and powerlessness. American Sociological Review, 1975, 40, 710-719.
- Patton, M. Q. Utilization-focused evaluation. Beverly Hills: Sage Publications, 1978.
- Phi Delta Kappa National Study Committee on Evaluation. Educational evaluation and decision making. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.
- Popham, W. J. Educational evaluation. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.
- Rokeach, M. The open and closed mind. New York: Basic Books, Inc., 1960.
- Rossi, P., Freeman, H., & Wright, S. Evaluation: A systematic approach. Beverly Hills: Sage Publications, 1979.
- Stecher, B. Constraints and encouragers to evaluation utilization. Paper presented at the American Educational Research Association Annual Meeting, Los Angeles, 1981.
- Triandis, H. C. Attitude and attitude change. New York: John Wiley & Sons, 1971.
- Weiss, C. H. Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, Inc., 1972.
- Worthen, B. R., & Sanders, J. R. (Eds.). Educational evaluation: Theory and practice. Worthington, Ohio: Charles A. Jones, 1973.

ACKNOWLEDGEMENTS

I am very grateful to all of the people who contributed to the successful completion of this research project.

Specifically, I wish to thank the members of my graduate advisory committee, each of whom made a unique contribution. Dr. Anton Netusil and Dr. Ross Engel were the deciding factors for my receiving cooperation from local school districts. When public school administrators learned that these two highly respected men supported my project, they willingly cooperated. Dr. W. Andrew Hunter's contemplative nature positively influenced the development of this dissertation. Dr. Charles Mulford offered positive comments and encouragement. Also, Dr. Mary E. Huba provided valuable personal and professional guidance, technical assistance, and the highest of standards. I am very fortunate to have had such a role model to influence, teach, and help me during the last three years. Finally, I am thankful to all five of these professors for taking the time to read this dissertation carefully and to offer constructive recommendations for its improvement.

Another individual to whom I am grateful is Dr. Richard Warren. His contribution at every stage of development, especially the data analysis stage, of this study was extremely valuable. He treated me as a colleague when I was only a graduate student; this had a great impression on me. In summary, I feel sorry for the students attending Iowa State University who never have the opportunity to work with him.

My family also deserves acknowledgement. I appreciate my husband, Walter's constant confidence in me and expectation that I would attain this goal. The challenge provided me by my son, Hudie, who developed into toddlerhood as I was completing this dissertation, allows me to enjoy my accomplishment even more. Finally, I want to give credit to my parents, Rosella and Joseph Ruffo, for their unconditional support and love.

APPENDIX A. STANDARDIZED DIRECTIONS FOR RESEARCH MATERIALS

1. Introduce myself or acknowledge the introduction given.
2. Read first page of materials to group.
3. Ask if there are any questions.
4. "On the bottom of this page, please write the title of the position you held during the 1980-81 school year."
5. Also, please list the state in which you work.*
6. Now, on the front and back of the next page are questions about your background. On Q4, if you work in a junior high school, check "other" and then indicate what grades are included in your building. On Q6, if you pick a percent not listed, put an "X" on the line where your response is and in () put the number. On Q7, please read the definition of program evaluation carefully. These questions do not pertain to personnel evaluation. Please answer these and then stop until I give you further directions.
7. On page 3 are 10 questions. The first four are multiple choice, the last six are multiple choice. Please try to answer each question. There is no penalty for guessing. When you finish, please wait for further directions.
8. The next two pages ask for your opinions. From these responses I will try to identify thought processes or decision making styles. I will read through the directions with you. Read instructions. When you complete these items, please wait for further instructions.
9. The next 8 or 9 pages are for you to read carefully. Note that some people don't have page 7. That's all right. It will take between 20 and 30 minutes for you to complete this task. You will need to refer to these materials as you complete the last section of the packet, so feel free to underline or note anything you think is of importance. Also, if you need to leave the room, or would like to get more coffee, you may do so. Take your time, and read carefully. When you are finished, if you need to visit with one of your colleagues, please leave the room so as not to disturb those still reading.
10. These last pages ask for reactions to the evaluation materials you have just read. When you all have completed these, you will have a chance to ask questions and react to this exercise.
11. Any questions or comments?

*For workshop sample only.

12. Thank you very much. Please know that I am very grateful to you for your cooperation. Early this fall, I will send a summary of my findings to (_____) (contact person). He will disseminate it to you.

APPENDIX B. RESEARCH MATERIALS

This study is a simulation which will ⁷⁹ assess the potential impact of an evaluation. The purpose of this research is to provide guidelines to evaluators so that they can learn how to make evaluations more useful.

You will be asked to answer questions about your background, about your opinions relating to certain statements, and about an evaluation report you will read. In addition, other materials will be presented for you to read: a resume of the evaluator, a description of the program being evaluated and the setting in which the evaluation occurs.

An attempt will be made to answer any questions you have providing that they do not interfere with the experimental treatment of the study. You are under no obligation to participate and may withdraw from the study at any time. To insure complete confidentiality, do not put your name on the following pages.

Completion of these materials will imply that you have given your informed consent to participate in this research project. Your cooperation is truly appreciated.

EDUCATIONAL AND PROFESSIONAL BACKGROUND

1. For how many years have you been a school administrator? ⁸⁰ _____ years

2. What is the enrollment in your district?

___ 250-1000

___ 1000-5000

___ 5000-10,000

___ 10,000-50,000

3. What is the enrollment in your building?

___ 70-150

___ 150-500

___ 500-1000

___ 1000-2500

4. What grades are included in your building?

___ 7-12

___ 9-12

___ 10-12

___ other

5. What is your highest degree?

___ B. S.

___ M. S.

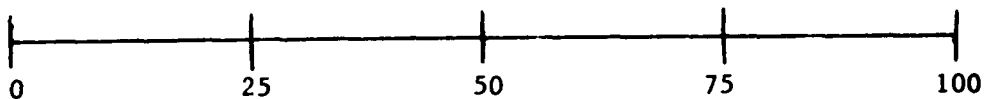
___ Specialist or M. S. +45

___ Ed. D. or Ph. D.

6. On a scale from 0 to 100, assess the impact that your opinion has on the final hiring, firing, and transfer decisions in your school.

My input:

has no impact is rarely considered is sometimes considered is always considered is the only opinion considered



7. Program evaluations are evaluations that assess educational activities which provide services on a continuing basis and often involve curricular offerings. Listed are possible ways you could have been involved in a program evaluation. Check ✓ if you have had such experience.

81

Have you ever:

- been directly responsible for a Title I evaluation at your school or at any other school
- been involved in a Title I evaluation at your school or at any other school
- been directly responsible for a State mandated program evaluation at your school or at any other school?
- been involved in a State mandated program evaluation at your school or at any other school?
- been directly responsible for a locally initiated program evaluation at your school or at any other school?
- been involved in a locally initiated program evaluation at your school or at any other school?
- planned and implemented a program evaluation at your school?
- requested funding for an evaluation at your school?
- suggested to the superintendent or board that an evaluation was needed for a program at your school?
- made a decision based on information obtained through an evaluation?
- worked with an evaluator by providing him/her with information or data about a program?
- served on a team or committee which evaluated a program?
- written a program evaluation report?
- read a program evaluation report?
- asked the DPI for help with a program evaluation?
- had a program at your school evaluated to see if it qualified as a State or Federally validated program to be disseminated (example, as through the National Diffusion Network)?
- attended a workshop on program evaluation?
- taken a graduate course in program evaluation?

Please answer each question selecting the correct answer, or filling the blank with the correct answer.

82

- _____ 1. The basic purpose of educational evaluation is to:
- test judgemental hypotheses.
 - improve decision making.
 - test the generalizability of formative data.
 - validate the findings of basic research in field settings.
- _____ 2. A basic principle of goal-free evaluation is that
- the evaluator should not know in advance the program goals.
 - the evaluator should not know in advance the decision that needs to be made about the program.
 - the evaluation design should not have goals.
 - the evaluation should be organized around behavioral objectives rather than goals.
- _____ 3. CIPP is an acronym for
- continuous improvement of programs through planning.
 - collaborative input into program planning.
 - context, input, process, and product.
 - continuity, input, planning, and process.
- _____ 4. The first step in the PDK Evaluation Model is
- Needs Assessment
 - Behavioral objectives
 - Audience Identification
 - Delphic Probe
- _____ 5. Who is commonly referred to as the 'Father of Educational Evaluation?
- _____ 6. An evaluation which assesses the merits of an already completed project or program is a(n) _____ evaluation.
- _____ 7. The system of measurement which compares an individual student's performance against the attainment of the objectives of the program rather than against the achievement of other learners is called a _____ system of measurement.
- _____ 8. The CSE(Center for Study of Evaluation) is at what University?
- _____ 9. One problem with using a certain kind of test in an evaluation is that there is often a substantial lack of congruence between what the test measures and what is stressed in a local curriculum. To what kind of test does this problem refer?
- _____ 10. An evaluation that provides information for improvement of a program that is still under development is a(n) _____ evaluation.

Directions: The following is a questionnaire of what the general public thinks and feels about a number of important social and personal questions. The best answer to each statement below is your personal opinion. The authors have tried to cover many different and opposing points of view; you may find yourself agreeing strongly with some of the statements, disagreeing just as strongly with others, and perhaps uncertain about others. Whether you agree or disagree with any statement, you can be sure that many people feel the same as you do.

After each statement, circle the number which corresponds to your personal opinion. The scale is as follows:

- | | |
|--------------------------|-----------------------------|
| +1: I agree a little | -1: I disagree a little |
| +2: I agree on the whole | -2: I disagree on the whole |
| +3: I agree very much | -3: I disagree very much |

RESPOND TO EACH STATEMENT.

- | | | | | | | |
|--|----|----|----|----|----|----|
| 1. There are two kinds of people in this world: those who are for the truth and those who are against the truth. | +3 | +2 | +1 | -1 | -2 | -3 |
| 2. It is only when a person devotes himself to an ideal or cause that life becomes meaningful. | +3 | +2 | +1 | -1 | -2 | -3 |
| 3. The United States and Russia have just about nothing in common. | +3 | +2 | +1 | -1 | -2 | -3 |
| 4. In the history of mankind there have probably been just a handful of really great thinkers. | +3 | +2 | +1 | -1 | -2 | -3 |
| 5. A group which tolerates too much differences of opinion among its own members cannot exist for long. | +3 | +2 | +1 | -1 | -2 | -3 |
| 6. There are a number of people I have come to hate because of the things they stand for. | +3 | +2 | +1 | -1 | -2 | -3 |
| 7. Once I get wound up in a heated discussion I just can't stop. | +3 | +2 | +1 | -1 | -2 | -3 |
| 8. Unfortunately, a good many people with whom I have discussed important social and moral problems don't really understand what's going on. | +3 | +2 | +1 | -1 | -2 | -3 |
| 9. In this complicated world of ours the only way we can know what's going on is to rely on leaders or experts who can be trusted. | +3 | +2 | +1 | -1 | -2 | -3 |

- 84
25. The present is all too often full of unhappiness.
It is only the future that counts. +3 +2 +1 -1 -2 -3
 26. A man who does not believe in some great cause
has not really lived. +3 +2 +1 -1 -2 -3
 27. My blood boils whenever a person stubbornly
refuses to admit he's wrong. +3 +2 +1 -1 -2 -3
 28. It is often desirable to reserve judgement about
what's going on until one has had a chance to
hear the opinions of those one respects. +3 +2 +1 -1 -2 -3
 29. It is only natural that a person would have a
much better acquaintance with ideas he believes
in than with ideas he opposes. +3 +2 +1 -1 -2 -3
 30. If given the chance I would do something of great
benefit to the world. +3 +2 +1 -1 -2 -3
 31. To compromise with our political opponents is
dangerous because it usually leads to the betrayal
of our own side. +3 +2 +1 -1 -2 -3
 32. In times like these, a person must be pretty
selfish if he considers primarily his own
happiness. +3 +2 +1 -1 -2 -3
 33. A person who thinks primarily of his own
happiness is beneath contempt. +3 +2 +1 -1 -2 -3
 34. Of all the different philosophies which exist in
this world there is probably only one which is
correct. +3 +2 +1 -1 -2 -3
 35. Most of the ideas which get printed nowadays
aren't worth the paper they are printed on. +3 +2 +1 -1 -2 -3
 36. Fundamentally, the world we live in is a pretty
lonesome place. +3 +2 +1 -1 -2 -3
 37. If a man is to accomplish his mission in life it
is sometimes necessary to gamble "all or nothing
at all". +3 +2 +1 -1 -2 -3
 38. It is only natural for a person to be rather
fearful of the future. +3 +2 +1 -1 -2 -3
 39. Most people just don't know what's good for them. +3 +2 +1 -1 -2 -3
 40. The highest form of government is a democracy
and the highest form of democracy is a government
run by those who are most intelligent. +3 +2 +1 -1 -2 -3

- 85
10. When it comes to differences of opinion in religion we must be careful not to compromise with those who believe differently from the way we do. +3 +2 +1 -1 -2 -3
 11. Even though freedom of speech for all groups is a worthwhile goal, it is unfortunately necessary to restrict the freedom of certain political groups. +3 +2 +1 -1 -2 -3
 12. The worst crime a person could commit is to attack publicly the people who believe in the same thing he does. +3 +2 +1 -1 -2 -3
 13. It is better to be a dead hero than to be a live coward. +3 +2 +1 -1 -2 -3
 14. In a discussion I often find it necessary to repeat myself several times to make sure I am being understood. +3 +2 +1 -1 -2 -3
 15. I'd like it if I could find someone who would tell me how to solve my personal problems. +3 +2 +1 -1 -2 -3
 16. While I don't like to admit this even to myself, my secret ambition is to become a great man, like Einstein, or Beethoven or Shakespeare. +3 +2 +1 -1 -2 -3
 17. Man on his own is a helpless and miserable creature. +3 +2 +1 -1 -2 -3
 18. Most people just don't give a "damn" for others. +3 +2 +1 -1 -2 -3
 19. In the long run the best way to live is to pick friends and associates whose tastes and beliefs are the same as one's own. +3 +2 +1 -1 -2 -3
 20. The main thing in life is for a person to want to do something important. +3 +2 +1 -1 -2 -3
 21. In times like these it is often necessary to be more on guard against ideas put out by people or groups in one's own camp than by those in the opposing camp. +3 +2 +1 -1 -2 -3
 22. In a heated discussion I generally become so absorbed in what I am going to say that I forget to listen to what the others are saying. +3 +2 +1 -1 -2 -3
 23. A person who gets enthusiastic about too many causes is likely to be a pretty "wishy-washy" sort of person. +3 +2 +1 -1 -2 -3
 24. There is so much to be done and so little time to do it in. +3 +2 +1 -1 -2 -3

86
A Comparison and Contrast of
Educational Research and Educational Evaluation*

Because educational research and educational evaluation are so frequently confused, and because the distinctions between these two activities are useful in better understanding the nature of systematic educational evaluation, the similarities and differences between the two will be explored.

There are many similarities between the activities of educational researchers and educational evaluators. They both engage in disciplined inquiry, use measurement devices and systematically analyze data. Also, they both describe their projects in reports.

Two distinctions between educational research and educational evaluation will be discussed: focus of the study and generalizability of the study's results.

FOCUS - Both researchers and evaluators are attempting to secure additional knowledge, but the use to which they wish to put this knowledge differs. Researchers want to draw conclusions and are interested in understanding relationships. Evaluators are more interested in decisions and want to guide someone's actions. Researchers focus on conclusions; evaluators focus on decisions.

GENERALIZABILITY - A pivotal difference between research and evaluation is the generalizability of the obtained results. An ideal research investigation would result in findings that could be generalized to a wide variety of comparable situations. The more generality that a researcher's findings have, the better the researcher likes it. Evaluation, on the contrary, pertains to a particular situation and what decisions to make about it. Usually, there is no intention of generalizing evaluation results to other situations.

*adapted from Popham, W. James. Educational Evaluation. Englewood Cliff, New Jersey:Prentice-Hall, Inc., 1975.

Information Treatment

The purposes of the new program, Applied Mathematics, were to raise level of basic math skills for Wilson High graduates, to provide the opportunity to learn practical, everyday uses of math, for all students regardless of ability, and to eliminate the stigma associated with enrolling in a 'remedial' course. Currently, there are four sections of Applied Math: two teachers teach two sections each. The course content has undergone some minor revisions since its development. The ten units which make up the course content are: Basic Operations, Fractions and Decimals, Percents and Ratio, Computer Literacy, Taxes, Insurance, Measurements I, Measurements II, Banking and Investments, and Logic. When this new program was implemented, the Remedial Math course was dropped from the list of available math courses.

The two staff members teaching Applied Math volunteered for the assignment. One of these, Mr. Roberts, had taught the only two sections of Remedial Math when it existed. He has worked at Wilson for 7 years; all of his teaching experience was gained here. Ms. Dennison, the other teacher, has had a total of 10 years of teaching experience; the last four have been at Wilson. Both teachers have Masters Degrees in Education, and describe themselves as highly organized but flexible. Mr. Roberts is involved in the 9th grade boys basketball and wrestling programs, as well as being advisor the the Honor Society. Ms. Dennison has no extra-curricular involvements. She is Secretary/Treasurer of the Faculty Association.

Major costs of the new course occurred during the first year. These included summer salaries for 3 weeks for 2 teachers who determined the exact content and curriculum. Ten major topics were selected, and an outline for a learning activity packet for each unit was written. Each packet was expanded during the first year of implementation. The teachers decided that no one textbook met their needs, so duplicating costs rather than textbook purchases are currently a major expense. Teachers drew materials from 5 textbooks. Seven copies of each of 5 books were purchased by the district. During the first year, a basic set of overhead transparencies was made for each unit. Drill and practice exercises for each unit were written by a student teacher, and are stored on the Intermediate Unit's computer for use as needed on Wilson's two computer terminals. These added no new direct costs as the district pays a fixed fee to the IU* for the use of the computer.

There are no prerequisites for the Applied Math course. There is a one credit math requirement for graduation. One credit is earned by the successful completion of two semesters of math during grades 9-12. Any student is permitted to enroll in this course. All other math offerings are either academically oriented, or specifically designed for a vocational program. Ninth graders who have a history of difficulty in math are encouraged to enroll. All other students are informed about the course through the student handbook and by other math teachers and can elect to take it.

General objectives have been written for each of the 10 units covered in the two semester course. All students must complete each unit to pass the course. However, the number and level of difficulty of the activities each student completes within each unit varies. The length of time spent on each unit is constant for the entire class. The materials used are stored in each classroom. One computer terminal is in Mr. Robert's classroom. The other is in the Counseling Suite.

*IU (Intermediate Unit) is similar to AEA(Area Education Agency)

Johnson City, Missouri is located in central Missouri and has a population of 25,000. The population and tax base are stable. The major industry is the Lilac B-B Gun Company. Like the national trend, school population here is declining. This drop necessitated the closing of one elementary school two years ago.

The school board consists of 6 members, elected for a three year term on a rotating basis. The superintendent has been at Johnson City for ten years and is considered successful by the community, board, faculty and administration. During his tenure, community concern was raised over four major events. First, due to declining enrollment, one elementary school was closed. Next, the threat of a teacher's strike was caused by a salary dispute. Third, a program to combat drug abuse was instituted in the high school. Last, there was a major district wide budget revision resulting from inflation. The public school district includes 5 elementary buildings(grades K-5), two middle schools(grades 6-8), and one high school(grades 9-12).

Wilson High School has 1200 students. The administration of the high school consists of a principal and two assistant principals. There are three counselors and 48 teachers. The faculty is organized in departments according to subject area. The average age of the staff is 31; the average length of teaching experience is 9 years; and the average salary for teachers is \$15,500. Over the past five years, an average of 300 students have graduated each year. Of that total, 70% pursue some type of post high school education. Twenty-five percent earn B.S. degrees; and 30% graduate from a junior college or a vocational/technical school.

In 1978, after a year of self-study, Wilson High was visited by an NCA team. The team highlighted the quality of the teaching staff, guidance services and the music program. The major problem observed by the team was the poor condition of the physical plant. The self-study reflected the same concern and a 5 year plan has been developed to make needed improvements.

A consultant, David Stoner, from Columbia, MO., was hired to evaluate the Applied Math program at Wilson High School. He was selected from a group of 20 applicants who were judged on the following qualifications: education, professional experience, current interests, publications and relevant evaluation experience. The math department budget has allocated \$2000 for the services of the consultant. The Applied Math program was initiated three years ago by the four math teachers at the school. When they decided to make a change from Remedial to Applied Mathematics, they planned to have an outside evaluator review the program during its third year. They hoped the evaluator would provide a judgement of overall quality, and if appropriate, suggestions for improvement of the Applied Math program. The principal is interested in knowing whether there is an improvement in the basic math skills of Wilson High graduates as a result of this program, and some ideas of the costs so that he can determine if the program is successful enough to continue or whether it should be changed or replaced. The principal will report and make recommendations to the Superintendent and School Board concerning this recently implemented math program.

DAVID STONER

This Resume was prepared specifically for the Evaluation Consultant position at Wilson High School.

EDUCATION:

- 1965 B.S. Secondary Education/Mathematics, Mankato State College, Minnesota
- 1968 M.S. Curriculum & Instruction, Iowa State University
Thesis: A Comparison of the Response of the General Public and Teacher Education Students When Asked to Describe a Good Teacher
- 1972 Ph.D. Curriculum & Instruction, Iowa State University
Dissertation: The Relationship Between Teachers' Attitudes toward NCA Evaluations and District Policy Changes

EXPERIENCE:

- 1965-1968 Junior High Math Teacher, Lincoln Junior High School, Shawnee Mission, KS.
- 1968-1972 Research Assistant, Research Institute for Studies in Education, Iowa State University
- 1972-1977 Assistant Professor, College of Education, University of Missouri/Columbia
- 1977-1980 Associate Professor, College of Education, University of Missouri/Columbia

CURRENT INTERESTS: Curriculum Development and Evaluation

PUBLICATIONS:

- "Teacher Attitudes Toward NCA Evaluation." Phi Delta Kappa, Vol. 72 (September, 1974),pp 36-39.
- "The Effect of Teacher Praise on Student Behavior." Journal of Educational Psychology, Vol. XV(February, 1975),pp 89-100.
- "Motivating Teachers to Evaluate the Curriculum." NASSP Bulletin, Vol. XX(January, 1976),pp 22-25.
- "An Effective Approach to Curriculum Planning." Educational Technology, Vol. 21(October, 1976),pp 5-12.
- "Subsequent Policy Change Resulting from NCA Evaluations." Educational Evaluation and Policy Analysis, Vol. 1(Fall, 1978),pp 40-46.
- "A System-Wide Attack on An Out-dated Curriculum." Educational Leadership, Vol. 40(July, 1979),pp 62-65.

RELEVANT EVALUATION EXPERIENCE: no directly applicable experience

Resume Treatment

DAVID STONER

90

This Resume was prepared specifically for the Evaluation Consultant position at Wilson High School.

EDUCATION:

- 1965 B.S. Secondary Education/Mathematics, Mankato State College, Minnesota
- 1968 M.S. Curriculum & Instruction, Iowa State University
Thesis: A Comparison of the Response of the General Public and Teacher Education Students When Asked to Describe a Good Teacher
- 1972 Ph.D. Curriculum & Instruction, Iowa State University
Dissertation: The Relationship Between Teachers' Attitudes toward NCA Evaluations and District Policy Changes

CERTIFICATIONS:

- 1969 Principal Certification, Iowa and Missouri
- 1975 Superintendent Certification, Iowa and Missouri

EXPERIENCE:

- 1965-1968 Junior High Math Teacher, Lincoln Junior High School, Shawnee Mission, Kansas
- 1968-1971 Assistant Principal, Lincoln Junior High School, Shawnee Mission, Kansas
- 1971-1974 Principal, Hickman High School, Columbia, Missouri
- 1974-1980 Assistant Superintendent for Curriculum and Instruction
Columbia Independent School District, Missouri

CURRENT INTERESTS: Curriculum Development and Evaluation

PUBLICATIONS: No Publications

RELEVANT EVALUATION EXPERIENCE: Consulting

- 1974-1978 Team Leader for NCA Evaluation in the Midwest, Conducted 10 Evaluations during that time
- 1976 member of the Missouri State Department of Public Instruction's Task Force which evaluated statewide graduation requirements in math and science
- 1977 Evaluator for the Gifted & Talented Math Program at the elementary level in St. Louis Public School District
- 1977 Evaluator for a Title IVC Grant at North High School in Springfield, MO. Program Title: Continuous Progress Math--Grades 1-8.
- 1978-1980 Chairman of the Consulting Team for Hannibal School District, Hannibal, MO., Evaluation of the Effectiveness of the Scope and Sequence of the Mathematics Curriculum Grades 1-12.
- 1978 Evaluator of the Elementary Math Curriculum, Wallington School District, MO.
- 1979 Evaluator of the Vocational Math Program at Johnston High School, Johnston, KS.

RELEVANT EVALUATION EXPERIENCE: Current Position

- 1975 Parental and Teacher Satisfaction with the Primary Math Program
- 1976-1979 Evaluation of the Effectiveness of the Scope and Sequence of the Social Studies Curriculum Grades 1-12
- 1977 Evaluation of the College Bound Math sequence
- 1979 Needs Assessment and Student Achievement in Math and Reading

Resume Treatment

A meeting was held with the evaluator, math teachers and principal of Wilson High soon after the evaluator was hired. At this meeting, David Stoner agreed that the final report would be given to the principal and math department on or before March 1. Also, Stoner presented the staff and administrator with a list of twenty potential evaluation questions he could study, and asked them to select the four most important. They selected these four:

1. How does this program fit into the entire K-12 math curriculum?
Are there still needs of students that are not being met?
2. Are students gaining competencies in practical uses of math?
3. Is the Applied Math course equipping low achievers with basic competencies?
4. Are there any important positive or negative outcomes of the program which were not anticipated by the developers?

In the report which follows, the process used to gather information and an answer to each question is presented.

February 20, 1981

Mr. Samuel McManus
Principal
Wilson High School
Johnson City, Missouri

Dear Sam,

For the past ten weeks I have been working with you and the math department staff evaluating the Applied Math Program at Wilson High School. During the first two weeks I learned about the course by reading materials and talking with you, Ms. Dennison, and Mr. Roberts. After this initial period, your committee selected four areas of concentration from a list of twenty possible areas. The group decided that I should concentrate my efforts on supplying answers to these four questions. The report which follows is an executive summary of the entire evaluation report document which will be forwarded to you on March 1. The final document will include all materials used, and all raw data analyzed.


Most of the findings fo this evaluation are positive. A few areas for improvement were identified and strategies to strengthen them are recommended. Based on these findings my overall recommendation is that the Applied Math course be continued and that modifications be made where indicated.

The costs of implementing the improvements relate to staff time needed to develop new and change existing materials. These changes are not major.

Finally, the purposes of the course shoud be reviewed. The four areas selected for my efforts did not include many of the purposes stated in the program description. The stated purposes are: to raise the level of basic math skills for Wilson High graduates, to provide the opportunity to learn practical, everyday uses of math, for all students regardless of ability, and to eliminate the stigma associated with enrölling in a 'remedial course'. Either the purpose statement should be updated or an effort should be made to see that the purposes are being accomplished.

I have enjoyed working with you and your fine staff. I commend you on your concern with quality education and encourage your efforts to evaluate your programs. I am available to elaborate on any findings whenever your group requests. You will be receiving the final report from me during the next ten days.

Sincerely,


David Stoner,
Evaluator

THE EXECUTIVE SUMMARY
of the
APPLIED MATH PROGRAM EVALUATION
at
WILSON HIGH SCHOOL
Johnson City, Missouri

How does this program fit into the entire K-12 math curriculum? Are there still needs of students that are not being met?

94

A vertical math curriculum committee was appointed in 1974 and charged with the task of creating a K-12 Scope and Sequence Chart that would be the visual organization of the Johnson City School District's mathematics program. In May, 1975, the completed project was reviewed by all of the district math teachers and adopted with plans to review it for major revisions in 1985.

Using this document as a guide, the 10 units of the Applied Math course were compared to the areas outlined in the Scope and Sequence Chart. This identified any areas mentioned in the Chart that were not covered in the outline of units in Applied Math, as well as identifying any extra areas covered in Applied Math not mentioned in the Chart. Graphing and Geometric Principles are two major areas described in the Scope and Sequence Chart that are not listed as units in Applied Math. The units on Computer Literacy, Investments and Logic are taught in Applied Math but not included in the K-12 Chart.

Because Applied Math is the final math course that some students at Wilson High School take, it is very important to assess the extent of the discrepancy between the course outline and the goals established at the district level as described in the Chart. Therefore, objectives of the 10 units were carefully studied to determine if content about Graphs, or Geometric Principles was included in the course. In the Measurement II unit, objectives pertaining to perimeter, area and volume are outlined. There was no mention of the study of angles, for example, which is a topic described in the Chart. Further, in the course objectives there were no objectives concerning the construction of graphs, or interpretation of data presented in graphic form.

Although the topics of Computer Literacy, Logic and the subtopic, Investments, are not required through the Scope and Sequence Chart, they were determined as useful by the teachers developing the Applied Math course. Therefore, perhaps the length of time spent on each unit could be modified to allow inclusion of the two lacking areas, graphs and geometry, so that none of these three 'extra' areas needs to be dropped. It is also possible that the graph and geometry topics could be incorporated into existing units.

Teacher's guides from several high school level Applied or Practical Math books were obtained. Chapter or Unit tests, as well as final tests from these guides were used as sources for items to be used on a test to assess the level of competence in the practical uses of math outlined in the Applied Math course. If items were unavailable from these sources, the evaluator constructed them. For example, no items for any Computer Literacy objectives could be found. Therefore, the evaluator wrote some. For the first step in building the test, the evaluator selected 25 items for each of the 10 units that he thought reflected the course objectives. That is a total of 250 items. Those same items were studied by the teachers and 40 of the items were eliminated. Although there were some course objectives without corresponding items, the teachers were comfortable with the final 210 items as those best measuring the most important objectives.

The population of students to be tested were the sophomores, juniors, and seniors at Wilson High who had completed the Applied Math course in either of the two previous years. The students were asked to come to the cafeteria during their study hall to take the test. It was explained that no names should be put on the test papers, as the students were not being evaluated by the test. However, to assist the evaluator in answering the next evaluation question, students were asked to indicate their semester grades and final grade for their 8th grade math class. Students who had no study hall or who chose not to participate were not included in the study. During one week in November, 150 students cooperated.

Because of time limitations all 210 items were not administered to all 150 students. Each student responded to 50 multiple choice items; 5 items each for the 10 units. However, in total, all 210 items were used. This method of testing is permissible because of the fact that it is the program being evaluated rather than the individual student.

Although the complete set of results is available, only those items to which less than 50% of the responses were correct are presented here. The math teachers chose 50% as the point at which they would be concerned. The items numbers, the unit to which each corresponds and the percent of students answering the item correctly are listed.

Item No.	Unit	% Correct
12	Insurance	48
14	Fractions & Decimals	47
17	Computer Literacy	35
22	Basic Operations	42
27	Basic Operations	39
28	Insurance	38
32	Percents % Ratios	46
33	Computer Literacy	47
39	Insurance	29
47	Measurements I	46
49	Taxes	37
52	Taxes	46
63	Taxes	40
67	Measurements II	43
68	Measurements II	45
70	Insurance	35
77	Banking & Investments	43
90	Basic Operations	47
99	Computer Literacy	25
100	Taxes	39
118	Logic	29
121	Percents & Ratios	35
176	Fractions & Decimals	22
187	Percents & Ratios	22
192	Logic	27
198	Percents & Ratios	47
200	Percents & Ratios	39
201	Taxes	29
204	Computer Literacy	38
206	Banking & Investments	37
209	Logic	42
210	Fractions & Decimals	41

Summary:

Unit	Number of Items
1. Basic Operations	3
2. Fractions & Decimals	3
3. Percents & Ratios	5
4. Computer Literacy	4
5. Taxes	5
6. Insurance	3
7. Measurements I	1
8. Measurements II	2
9. Banking & Investments	2
10. Logic	3

Proportionately more items were missed in each of the first three units than in each of the last seven. The last seven are more directly concerned with applying math principles to everyday living. A total of 32 items out of the total of 210 were responded to with a rate less than 50% correct. From these results, students are acquiring greater competence in the Measurement, Banking and Investments and Logic areas than they are in the areas of Percents and Ratios and Taxes.

Is the Applied Math course equipping low achievers with basic competencies?

97

Low achievers were defined as students who received the grade of 'C' or 'D' as the final grade, or an 'F' as either semester grade in their 8th grade math course. Those low achieving students who are now in 10th, 11th or 12th grade who have completed the Applied Math course at Wilson High School were studied. The units of the Applied Math course which were defined as basic competency areas are: Basic Operations, Fractions and Decimals, Percents & Ratios and Measurement I.

The results of the test used to answer the second evaluation question about competencies in practical areas of math were reanalyzed and used to answer this question about low achievers and basic competencies. The responses of low achievers on the questions dealing with the four basic units were selected. The results of item numbers, corresponding units, and percent correct when that percent is less than fifty, are listed below. It was assumed that these low achievers enter the applied math class with lower basic competencies than average or high achievers do.

A total of 95 low achievers were identified.

Item No.	Unit	% Correct	
14	Fractions & Decimals	42	
19	Fractions & Decimals	40	
22	Basic Operations	35	
25	Basic Operations	42	
27	Basic Operations	33	
29	Basic Operations	47	
32	Percents & Ratios	40	
47	Measurements I	40	Summary:
53	Percents & Ratios	47	
55	Fractions & Decimals	45	Unit
59	Fractions & Decimals	47	
62	Percents & Ratios	45	Basic Operations
67	Measurements I	47	Fractions & Decimals
83	Fractions & Decimals	48	Percents & Ratios
90	Basic Operations	40	Measurements I
93	Measurements I	48	
121	Percents & Ratios	30	
150	Basic Operations	45	
176	Fractions & Percents	19	
186	Percents & Ratios	46	
187	Percents & Ratios	19	
198	Percents & Ratios	42	
200	Percents & Ratios	30	
210	Fractions & Decimals	35	

Are there any important positive or negative outcomes of the program not anticipated by the developers?

Two negative outcomes of the course have resulted from the large amount of organization needed to successfully implement the program. Through observation of classes and discussions with the two teachers instructing the course the evaluator concludes that there is a problem with teachers using class time efficiently. Because each student may be working on a different task, teachers can assist only individually or in small groups. Therefore, some students must wait for the teachers help and for a while be off task. Also, some above average students get finished with a unit up to one week before the slower students, despite the fact that the tasks in each unit for each student are individualized. Therefore, the above average student may be in class for five days with no work to do in mathematics.

Because of the above reasons, it is difficult to get a substitute teacher who can successfully manage the class. Some qualified substitutes have refused to replace either Mr. Roberts or Ms. Dennison. This information was given to the evaluator by the principal when discussing the unanticipated outcomes of the Applied Math program.

Two encouraging outcomes are the positive attitudes of the math department staff, and the improved quality of tests used in the program. All four math teachers share ideas and suggestions about the course and all appear genuinely committed to its success. Another factor contributing to the positive attitude of the staff is the recognition the course has received. On three different occasions, teachers from area schools have requested permission to visit the Applied Math classes, see the materials and learn about its organization. Some have suggested that the teachers at Wilson investigate having the DPI evaluate the course to see if it meets the criteria to have it disseminated as a state validated program.

The other positive outcome is that the teachers have continually collected data on the post tests for each unit, allowing them to produce high quality criterion-referenced exams for the course, and to modify the course content when needed.

RECOMMENDATIONS

99

Recommendation I: The Applied Math course should be modified so that it includes objectives pertaining to (1) geometry and (2) graphing, using the district's Scope and Sequence as a guide.

- A. The Measurement I and Measurement II units should be changed to include geometry objectives, such as, differentiating right, acute, and obtuse angles.
- B. A unit covering graphing objectives needs to be developed.
- C. To make time in the schedule for these units the Computer Literacy, and Logic units should be made optional for students who complete all other units.

Recommendation II: Low achievers need more work in the areas of Basic Operations, Fractions & Decimals, and Percents & Ratios

- A. Low achievers could be identified by their previous math achievement so that they can receive additional help during the study of these three units.
- B. This help could be in small group instruction by the teacher, or by one to one work with the teacher, or with a more advanced student.
- C. The computer drill and practice exercises should be emphasized for these students.

Recommendation III: The class organization needs flexibility to accommodate those students who complete units ahead of schedule.

- A. These students might assist the teacher as tutors.
- B. The optional units of Computer Literacy, or Logic can be utilized at this time.
- C. Students could select independent projects to work on during their extra class time. Example: How math skills are used in different career fields.

Recommendation IV: One or two substitute teachers who can manage these classes and who are trained as math teachers should be identified. If either Applied Math teacher must be absent, Johnson High should request the services of one on those substitutes.

The following questions relate to the materials you have just read. Considering the setting in which this evaluation took place, please respond to these items to the best of your ability.

100

1. Please rate the person conducting this evaluation on these twelve traits.

- (1) means that a minimal level of the quality was evident
- (5) means that a maximum level of the quality was evident
- (N) means that you have no opinion

Please circle your response.

This evaluator was:	minimum					maximum	no opinion
a. thorough	1	2	3	4	5	N	
b. unqualified	1	2	3	4	5	N	
c. knowledgeable about testing and grading	1	2	3	4	5	N	
d. unaware of school needs	1	2	3	4	5	N	
e. illogical	1	2	3	4	5	N	
f. practical	1	2	3	4	5	N	
g. believable	1	2	3	4	5	N	
h. convincing	1	2	3	4	5	N	
i. subjective	1	2	3	4	5	N	
j. fair	1	2	3	4	5	N	
k. expert	1	2	3	4	5	N	
l. technically competent	1	2	3	4	5	N	

2. Would you recommend the services of this evaluator to a colleague that you knew and respected?

- yes
- maybe
- no

3. Below are some applications of using evaluation information. In the situation at Wilson High and the Applied Math Program Evaluation, some of these may be appropriate; others may not. As principal of Wilson High, for which of the following purposes would you feel comfortable using this evaluation information? That is, which of the following would be legitimate uses?

101

Check all that apply. Under 'other' list any possible uses not mentioned.

- a. change or modify practices and procedures
- b. confirm what you thought was the situation when you first read about the program
- c. continue or discontinue the program
- d. allocate resources differently
- e. make personnel changes
- f. gain recognition or support for the program
- g. affect people's perceptions about the desirability of the program's goals
- h. affect or influence long range planning
- i. provide information for papers or articles about the program

OTHER

j.

The following set of questions will be repeated 5 times. One time for each of the four evaluation questions discussed by the evaluator in the report, and one time for the evaluation in total. Select the ratings as follows:

- 5 means very useful
- 4 means mostly useful
- 3 means midly useful
- 2 means mostly useless
- 1 means completely useless

Question #1: How does this program fit into the entire K-12 math curriculum? Are there still needs of students that are not being met?

Considering the information provided in response to this evaluation question,

a. How would you rate the usefulness of the evaluation process?

1 2 3 4 5

b. How would you rate the usefulness of the information itself?

1 2 3 4 5

21.

Question #2: Are students gaining competencies in practical uses of math

102

a. How would you rate the usefulness of the evaluation process?

1 2 3 4 5

b. How would you rate the usefulness of the information itself?

1 2 3 4 5

Question #3: Is the applied math course equipping low achievers with basic competencies?

a. How would you rate the usefulness of the evaluation process?

1 2 3 4 5

b. How would you rate the usefulness of the information itself?

1 2 3 4 5

Question #4: Are there any important positive or negative outcomes of the program which were not anticipated by the developers?

a. How would you rate the usefulness of the evaluation process?

1 2 3 4 5

b. How would you rate the usefulness of the information itself?

1 2 3 4 5

Total evaluation

Considering the entire evaluation as described in the materials you have just read, please answer these three last questions.

a. How would you rate the usefulness of the evaluation process?

1 2 3 4 5

b. How would you rate the usefulness of the information itself?

1 2 3 4 5

c. Thinking of the overall impact that the evaluation could have, how would you rate its usefulness?

1 2 3 4 5

**APPENDIX C. PATH COEFFICIENTS AND RESIDUALS
FOR ORIGINAL MODEL (FIGURE 4)**

LISREL IV Output for Original Model (Figure 4)

BETA - Regression coefficients for dependent variables:

	USES	EQ	USEFUL	OAIMPACT
USES				
EQ				
USEFUL				
OAIMPACT	.2549*	-.0553*	-.7485**	

GAMMA - Regression coefficients for predictor variables:

	RTRT	ROKEACH	KNOW	YRS	EVALEXP	DMR	ITRT
USES			.8417*	.0292	.0058		.1358
EQ	1.6249			-.2611	-.1015	-.0805	.3815
USEFUL		.0215	.7924		-.0279	-.0133	
OAIMPACT							

*Significant at .05 level.

**Significant at .01 level.

Table A

Residuals for Original Model

	USES	EQ	UP	UI	OAIMPACT	RTRT	ROKEACH
USES	0.00						
EQ	6.42	-0.03					
UP	1.62	10.71	0.00				
UI	1.85	11.48	-0.02	0.00			
OAIMPACT	1.46	6.39	0.29	0.23	0.23		
RTRT	0.09	0.01	0.28	0.17	0.06	0.0	
ROKEACH	-3.99	-42.29	-0.23	-4.08	3.78	0.07	-0.10
KNOW	-0.10	-0.22	-0.37	-0.28	-0.18	-0.04	1.14
YRS	0.39	0.07	0.85	0.45	1.49	-0.01	-0.14
EVALEXP	0.26	0.17	-0.13	0.18	0.27	0.00	-0.32
BENR	0.21	-0.08	-0.16	-0.06	-0.15	-0.05	0.21
DENR	0.03	-1.76	-0.30	-0.51	-0.29	-0.08	3.01
POS	0.10	0.18	0.06	0.03	0.09	0.01	0.03
IMPACT	3.01	19.39	4.01	0.63	6.89	-0.01	-45.61
ITRT	0.01	-0.01	-0.02	0.01	0.02	0.00	0.01

KNOW	YRS	EVALEXP	BENR	DENR	POS	IMPACT	ITRT
0.03							
-0.10	-0.04						
-0.03	0.00	0.01					
-0.01	0.71	1.36	0.00				
-0.11	1.57	0.65	0.13	0.00			
0.00	-0.16	0.01	0.03	-0.01	0.00		
-2.50	-0.06	-23.17	-1.65	0.99	-0.33	-0.50	
0.00	0.00	-0.01	-0.02	-0.02	0.00	-0.55	0.00

**APPENDIX D. PATH COEFFICIENTS AND RESIDUALS
FOR REVISION 1 (FIGURE 5)**

LISREL IV Output for Revision 1 (Figure 5)

BETA - Regression coefficients for dependent variables:

	EQ	USEFUL	USES	OAIMPACT
EQ				
USEFUL				
USES	-.1328*	-.0677		
OAIMPACT	-.0109	-.7470**		

GAMMA - Regression coefficients for predictor variables:

	RTRT	ROKEACH	KNOW	YRS	EVALEXP	DMR	ITRT
EQ	-2.5223	.2829		-.2535	27.9022	8.0352	112.1
USEFUL		.0629	.4355		6.0892	1.7495	24.5911
USES			.6120*	.0648*	.0129		.0083
OAIMPACT							

*Significant at .05 level.

**Significant at .01 level.

Table B

Residuals for Revision 1

	EQ	UP	UI	USES	OAIMPACT	RTRT
EQ	-0.09					
UP	-0.39	-0.01				
UI	-0.01	-0.04	-0.01			
USES	0.00	0.02	0.19	0.00		
OAIMPACT	0.14	0.02	-0.01	-0.33	-0.01	
RTRT	-0.02	0.08	-0.02	0.04	-0.05	0.0
ROKEACH	-0.92	0.22	-3.68	0.99	4.68	0.01
KNOW	0.00	-0.02	0.08	0.01	-0.06	-0.01
YRS	-0.05	0.29	-0.12	0.01	0.44	-0.01
EVALEXP	0.22	-0.32	-0.01	0.05	-0.31	-0.05
POS	-0.06	-0.10	-0.14	-0.01	0.00	0.00
IMPACT	5.38	-0.92	-4.51	1.27	4.44	-0.83
ITRT	0.00	-0.03	0.00	0.00	0.04	0.00

ROKEACH	KNOW	YRS	EVALEXP	POS	IMPACT	ITRT
-0.17						
-0.17	-0.01					
0.16	-0.01	0.01				
0.01	-0.33	-1.50	0.84			
0.79	-0.09	-0.89	0.09	0.00		
-12.86	-3.40	-6.08	6.99	2.19	0.70	
0.01	0.00	-0.01	-0.01	0.03	-0.67	0.00

**APPENDIX E. PATH COEFFICIENTS AND RESIDUALS
FOR REVISION 2 (FIGURE 6)**

LISREL IV Output for Revision 2 (Figure 6)

BETA - Regression coefficients for dependent variables:

	EQ	USEFUL	USES	OAIMPACT
EQ				
USEFUL				
USES	-.1472**			
OAIMPACT		-.7699**		

GAMMA - Regression coefficients for predictor variables:

	RTRT	ROKEACH	KNOW	YRS	EVALEXP	DMR	ITRT
EQ				-.2427			
USEFUL			.5604				
USES			.6191*	.0710*			
OAIMPACT							

*Significant at .05 level.

**Significant at .01 level.

Table C

Residuals for Revision 2

	EQ	UP	UI	USES	OAIMPACT	RTRT
EQ	0.00					
UP	11.42	0.00				
UI	12.22	-0.01	0.00			
USES	0.24	1.94	2.19	0.06		
OAIMPACT	9.49	0.01	0.00	1.17	0.00	
RTRT	0.45	0.32	0.21	0.11	0.12	0.0
ROKEACH	-45.50	-9.49	-13.75	-5.36	-3.10	0.00
KNOW	0.08	-0.10	0.00	0.04	-0.14	-0.02
YRS	0.50	0.31	-0.10	0.11	0.37	0.00
EVALEXP	-0.77	-0.48	-0.17	0.40	-0.49	0.00
POS	-0.09	-0.05	-0.09	0.03	0.04	0.01
IMPACT	5.58	1.71	-1.81	3.49	6.64	-0.22
ITRT	0.24	0.05	0.09	0.02	0.11	0.00

ROKEACH	KNOW	YRS	EVALEXP	POS	IMPACT	ITRY
0.02						
0.45	-0.01					
-0.71	-0.01	0.01				
-0.02	-0.08	-0.05	0.03			
0.79	-0.02	-0.50	0.08	0.00		
-15.50	-0.42	9.47	-1.68	0.00	0.02	
0.00	-0.01	0.00	-0.01	0.03	-0.67	0.00

**APPENDIX F. PATH COEFFICIENTS AND RESIDUALS
FOR REVISION 3 (FIGURE 7)**

LISREL IV Output for Revision 3 (Figure 7)

BETA - Regression coefficients for dependent variables:

	EQ	USEFUL	USES	OAIMPACT
EQ				
USEFUL				
USES	-.1111*	-.1497*		
OAIMPACT	-.0356	-.6985**		

GAMMA - Regression coefficients for predictor variables:

	RTRT	ROKEACH	KNOW	YRS	EVALEXP	DMR	ITKT
EQ		.1458*		-.1317			
USEFUL	.8644	.0345*	.3465				
USES			.5088*	.0609*			
OAIMPACT							

*Significant at .05 level.

**Significant at .01 level.

Table D

Residuals for Revision 3

	EQ	UP	UI	USES	OAIMPACT	RTRT
EQ	-0.12					
UP	9.21	0.00				
UI	9.94	-0.05	0.00			
USES	1.32	0.98	1.20	0.25		
OAIMPACT	6.28	0.33	0.32	0.33	0.35	
RTRT	0.39	0.07	-0.03	0.07	-0.05	0.0
ROKEACH	3.68	2.86	-1.02	1.67	7.40	0.12
KNOW	-0.22	0.00	0.10	0.04	-0.04	0.00
YRS	0.14	1.68	1.30	0.28	1.63	0.00
EVALEXP	-1.24	0.16	0.50	0.47	0.08	0.00
POS	-0.03	-0.09	-0.13	0.03	0.00	0.01
IMPACT	8.77	-0.69	-4.28	3.36	4.34	-0.22
ITRT	0.35	0.06	0.10	0.04	0.11	0.00

ROKEACH	KNOW	YRS	EVALEXP	POS	IMPACT	ITRT
2.94						
-0.18	-0.02					
0.94	-0.07	0.00				
-0.15	-0.11	-0.05	0.03			
0.82	-0.02	-0.50	0.08	0.00		
-13.90	-0.16	9.50	-1.61	0.00	0.02	
0.17	0.00	0.00	-0.01	0.03	-0.66	0.00
