BUILDING A FOUNDATION FOR
EVALUATION OF INSTRUCTION IN HIGHER EDUCATION
AND CONTINUING EDUCATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Ronald Jerry Adams

* * * * *

The Ohio State University
1973

Reading Committee:
Dr. Virgil Blanke  
Dr. William E. Loadman
Dr. James K. Luncan

Approved By

Advisor
Faculty of Educational Development
BUILDING A FOUNDATION FOR EVALUATION OF INSTRUCTION IN HIGHER EDUCATION AND CONTINUING EDUCATION

Abstract

The objectives of this exploratory study were sequentially as follows:

1) To develop a descriptive list of resources for use with evaluation of instruction;

2) To develop an evaluation of instruction kit;

3) To develop an evaluation of instruction workshop;

4) To evaluate the impact of the Evaluation of Instruction workshop;

5) To demonstrate the utility of the Evaluation of Instruction kit through its application in evaluating the Evaluation of Instruction workshop;

6) To assess use of the Evaluation of Instruction kit resulting from disseminating it without training potential users;

7) To assess the interim validity of the Evaluation of Instruction kit; and

8) To demonstrate the application of a potentially transferable method of interim validation of an educational product.
The focus of the study was on objective #7, to assess the interim validity (utility) of the Evaluation of Instruction kit. The prototype kit was considered valid partially to the extent that it was perceived as useful and was actually used by either trained or untrained potential users. Perceived utility of the kit was rated by potential users and by content experts. Actual use of the kit was assessed through telephone interviews with eight trained and six untrained recipients of the kit. The kit's utility was also demonstrated through the application of its contents to the evaluation of workshop training for a group of potential kit users. In addition, projected utility of the kit was rated by experts. Interim validity therefore is defined as consisting of a) actual use, b) perceived utility, c) projected utility, and d) demonstrated utility.

The conceptual framework of the kit incorporates components of the most prominent evaluation models. The ingredients in the kit were designed to encourage application of the framework of the kit to evaluation of instruction. Logically the Evaluation of Instruction kit has great potential value if instructors use it for improvement of instruction. The study is somewhat unique by focusing on the usefulness--rather than the revision--of an instructional product (the kit) during early
development. The study is also a demonstration of how an interim validation study can be performed.

Results from the study were encouraging. The first three objectives of the study were met by the creation of the Evaluation of Instruction Kit and by conducting the Evaluation of Instruction Workshop. The contents of the kit were used to evaluate the workshop, thereby meeting objectives four and five. A potentially transferable method of performing interim validation studies was generated to assess the interim validity of the kit, thereby meeting objectives seven and eight. The kit was disseminated to potential users without the workshop training, thereby meeting objective six. Although the kit was used by none of the six who received it without training, the kit was used by six out of the eight who were trained. Most indicators suggested that the training was a success and the kit was useful when accompanied with training.
ACKNOWLEDGEMENTS

The writer wishes to express thanks to Dr. Virgil Blanke for his support and assistance throughout the course of the writer's graduate work. To Dr. William Loadman, who helped conceptualize the study, who worked many hours to help create resources for the study, who defended the study before friend and foe, and who let his supper get cold so he could lend morale and technical support, the writer is certainly indebted. And to Dr. James K. Duncan, who gave suggestions and asked questions that assisted in focusing the study without discouraging the efforts of the writer, thanks is also extended.

The many kindnesses, words of cheer and unusually hard work of Ms. Carolyn Moss and Ms. Nancy Rutter were essential; the study would not have been possible without their special touch. Ms. Randy Kaplan, who typed the entire dissertation and who led the writer by the hand during the final hours of work must be given thanks for her unique and essential contribution.

Ms. Unhai Ahn, Ms. Barbara Janko, and Mr. Gary Milczarek provided excellent assistance in implementing the study. Ms. Helen Armstrong worked diligently to collect and pull together information regarding the study and then made a
final judgement regarding its worth. Thank you, Helen.

The Ohio Department of Mental Health and Mental
Retardation, Division of Mental Retardation, financed
a major portion of this study. Nisonger Center supplied
major support, also. The author readily acknowledges his
debt to both.

And to Katie Bones, the writer's wife, who gave
constant words of encouragement, technical assistance,
and major insights, the writer must give a free pass to
Disneyland.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>VITA</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF DEFINITIONS</td>
<td>xiii</td>
</tr>
</tbody>
</table>

**Chapter**

I. INTRODUCTION ........................................ 15  
   A. Preface  
      1. Why Did I Do It?  
      2. What I Did  
   B. Overview of Evaluation of Instruction Kit  
   C. Overview of Evaluation of Instruction Workshop  

II. REVIEW OF LITERATURE ............................... 39  
   A. Need for Kit and Validation  
   B. Alternatives to Evaluation Kit  
      1. Alternative evaluation frameworks  
      2. Complementary guides  
         a. Training packages  
         b. Guides for an overview of evaluation  
         c. Guides for performing evaluation  
      Alternative evaluation instruments  
      a. Compendia of listings and reviews  
      b. Compendia of evaluation instruments  
      Complementary instrument construction guides  
      a. Guides to constructing objectives  
      b. Guides for evaluation of students  
      c. Guides for student evaluation of instruction  
   3. Review of Alternatives Summation  

III. PROCEDURES ......................................... 97  
   A. General Statement  
   B. Participants for Study  
   C. Interim Validation Overview  
   D. Data Collection Framework for Evaluation of Workshop  
   E. Chronology of Assessment Activities  
   F. Threats to Validity of Design of Study  
   G. Reliability and Validity of Measures  
   H. Data Analysis By the Kit Developer  

viii
VITA

April 16, 1945. . .  Born - Jackson, Mississippi

1967. . . . . . .   B.A., University of Texas at Arlington
                 Arlington, Texas

1968-1970 . . . .  Training Specialist, Austin State
                 School, Austin, Texas

1970. . . . . . .  Completed coursework for M.A.
                 University of Texas, Austin, Texas

1971. . . . . . .  Consultant, The Research and Evaluation
                 Department of Cincinnati Public Schools,
                 Cincinnati, Ohio

1971. . . . . . .  Consultant, The Center for Vocational
                 and Technical Education, Columbus, Ohio

1971-1972 . . . .  Research Associate, Evaluation Center
                 The Ohio State University, Columbus,
                 Ohio

1972. . . . . . .  Evaluator, Center for Research on
                 Utilization of Scientific Knowledge,
                 University of Michigan, Ann Arbor,
                 Michigan

1972-1973 . . . .  Research Associate, Nisonger Center for
                 Mental Retardation and Developmental
                 Disabilities, The Ohio State University,
                 Columbus, Ohio

PUBLICATIONS

"Group Play Activities." Austin, Texas: Texas Department

FIELD OF STUDY

Major Field: Educational Evaluation

x
LIST OF ILLUSTRATIONS

Figure
1. Sanders and Cunningham Framework for Formative Evaluation Activity ............................................. 47
2. Stake Framework ...................................................... 49
3. Stake Framework (adapted by Gish) ..................... 50
4. Hammond Framework ......... ........................................ 52
5. McGuire Framework .............................................. 55
6. Lesson Function Chart ............................................. 68
7. Background Comparisons ........................................... 99
8. Interim Validation Flow Chart .................................... 102
9. Evaluation of Instruction
   Interim Validation Flow Chart ................................... 104
10. Data Collection Framework for
    Assessing Kit Utility (Interim Validity) ......................... 105
11. Data Collection Framework for Workshop ..................... 108
12. Results By Objectives, ........................................... 131
LIST OF TABLES

Table                                                                 Page
1. Pre-Post Differences of Workshop Participants on Measures of Evaluation Skill, Knowledge, Achievement, and Opinion .................. 136
2. Differences Among Four Administrations of the Achievement Instrument ........... 138
LIST OF DEFINITIONS*

Several terms which may need clarification are:

CRITERION-REFERENCED INSTRUMENT: Assessment device which uses specified performance or objectives as a standard against which data are collected; it is developed to diagnose, prescribe, and assess mastery of objectives rather than to predict and compare behaviors.

ENTRY BEHAVIOR: Performance in affective, cognitive, or psychomotor domain before instruction.

(EDUCATIONAL) EVALUATION: A methodological activity that includes comparing, focusing, and describing and that has foci including decisions, objectives, and products.
   a. Evaluation Design: Data collection framework, at least for the purposes of this paper. More generally it includes plans for delineating, obtaining, and providing information.
   b. Formative Evaluation: Methodological activity focused on delineating, obtaining, and providing information on a product or process during development and generally for purposes of revision.
   c. Impact Evaluation: Methodological activity focused on effectiveness and occurring at the end of instruction; information is collected as a basis for judgement of worth. For purposes of this study, the worth of the kit is based primarily on its utility.

EXIT BEHAVIOR: Performance in affective, cognitive, or psychomotor domain after instruction.

EXPERIMENT: A type of research activity; it includes: randomization in selection and assignment of subjects, presence of a control group, and manipulation of a treatment.

GOAL: Global objective.

INSTRUCTION: Processes and products focused on systematic affective, cognitive, or psychomotor change in specified voluntary recipients.

*Definitions were derived from the writer's understanding of the terms from many sources.
INTERNAL CONSISTENCY: Intercorrelation among items on an assessment instrument; one measure of test reliability.

NEEDS ASSESSMENT: Systematic collection of information to describe relationship between intents and the actual.

PERFORMANCE OBJECTIVE: Intended outcome stated in measurable form and with criteria for success specified.

RESEARCH: A methodological activity focused on building new knowledge or testing knowledge already presented.

STATISTICIAN: A specialist in data analysis.

VALIDITY: The degree to which an assessment instrument measures what it purports to measure (for the field of measurement). Validation of an organizational framework, which the kit in this study represents, means assessment of utility, at least for the purposes of this study.

a. Content Validity: The degree to which an assessment instrument samples the specified universe of content; also the degree to which experts judge a product (a kit, for example) to have intrinsic worth.

b. Face Validity: The degree to which an assessment instrument appears to measure what it is intended to measure. For the purposes of this study, face validity is also the degree to which participants in the study perceive the Evaluation of Instruction kit as useful to them.

c. Interim Validity: The degree to which a product that is under development is assessed as useful or potentially useful. For the purposes of this study, interim validity refers to a) actual use, b) perceived utility, c) projected utility (in whole), d) projected utility (in parts), and e) demonstrated utility.

d. Predictive Validity: The relation of results on an assessment instrument to measures of performance at some future time.
Chapter I

INTRODUCTION

Preface

Why Did I Do It?

1. In the Beginning...

"Why pursue the field of evaluation?" the man said as I handed him the dissertation.

"How far back do you want to go?" I replied.

"Go back to the beginning."

I thought back several years as I began to speak. "We were in a desert cabin, growing our garden, and asking ourselves how we might best perform service.

We had to decide which arena to enter: Ecology? Religion? Politics? Education? or Philosophy?

Education seemed to be the most pervasive area. Almost everyone receives formal instruction.

After deciding to perform service in the field of education we needed to decide how to get at the core of education--the place where it begins and develops. We decided that formal education is based upon the values of the educators. Then we decided that educators express their values through the criteria they use in
making decisions. Educational evaluation involves the process of making criteria clear and operational. If the evaluator helps the educator clarify his criteria, if the evaluator helps the educator focus attention on important criteria, and if the evaluator increases the visibility of the educators' progress in reaching important criteria, then the vision of educators should become clearer. Then eventually we may all develop a clearer vision.

I wanted to learn and disseminate the methods of the evaluator. After choosing evaluation as the discipline within education, I needed to choose a focus of inquiry within evaluation. I chose evaluation of instruction.

Instruction seems to be a highly personal act -- the instructor performs much of his own evaluation, sets many of his instructional criteria, and makes many decisions about what he will continue. The most powerful focus of inquiry in evaluation seemed to involve helping the instructor get more useable and wholistic feedback on his own instruction." As I ended my soliloquy, the man began to ask "rapid-fire" questions.

2. The Focus

"But people are changed most easily when they are young. Why did you focus on instructors of higher
education and continuing education? Why didn't you work on elementary school teachers?"

"How do you influence children except through their teachers?" I replied. "How do you influence teachers except through their teachers? And how do you influence the teachers of teachers except by giving them a set of tools they will use voluntarily?"

The man had the dissertation open now and he was leafing through the Appendices. He pointed at the Appendices and asked, "What does all of this stuff have to do with what you just said?"

"Part of what you see in the Appendices is how I looked at whether some teachers of teachers would use a set of tools to improve their feedback. Hopefully, as they use more systematic feedback, the basis they use to make decisions will become clearer and improve."

"So why did you attempt to assess the interim validity, utility, I believe you call it, of a prototype kit?"

"Like I said, one of the most helpful things I thought I could do was to see whether teachers would actually use a set of tools, a kit, if it were available; the potential impact of the kit was large."

"Why compare two modes of dissemination?"
"If instructors found the kit useful without the preparation of a workshop, the potential audience for the kit may have expanded. The expense of the workshop and the lack of workshop teachers will limit the audience. So you see, the two modes of dissemination needed to be compared."

"Why are you examining this particular kit? You once told me that it isn't even in final form yet." A good question.

"Hmm. Well, it's the only set of tools of this kind that I know of. The conceptual framework of the kit incorporates components of the most prominent evaluation models; it would logically be of great value. The ingredients in the kit encourage the application of the model. I hope to find a lot of resources for the kit if it seems to be useful. Full development, validation, and dissemination can be very expensive. Now is the time to find out how acceptable it is likely to be. Developers need to find if they have users before a product is completed and huge resources have been expended."

3. More on Interim Validation

"Why do you call 'potential utility' by 'interim validity' rather than just plain 'potential utility'?"

"Most people would mistakenly think that they know what 'potential utility' means in this study. Interim
validity in this study refers to gaining a wide range of information on usefulness of an educational product during early development. It means more than asking people if the product 'looks useful!':"

"How else do you find potential usefulness?"

"You could make full use of the product and see if it holds up. You could mail your product out and check if people then begin using it. Perhaps you could send the product out and see if people are able to work with it. You could send it out and have experts judge the components. You could send it out to see if experts project that it will be helpful for the intended users."

"Why go to all of that trouble?"

"You will be testing an unfinished product so your information on the finished product will necessarily be limited. Checking in different ways is like taking several readings on a weak wind to be more sure of its direction. If you get positive responses from several sources you can get a clearer indication on the potential usefulness of the finished product. Perhaps no one will be able to use your product as it is, but perhaps it can be used if the user is trained first; perhaps potential users won't think the product looks useful to them, but perhaps experts will see strong potential in the
product under development; perhaps you won't be very confident about the product's utility until you have given it a grueling trial yourself to see how it can handle rough use. With a weak wind you need more than one reading."

4. What is Evaluative Research?

"Can I learn something important if I read your study and find that the particular kit you studied isn't useful?"

"If the particular kit does not prove useful, the demonstrations of evaluation methodology should still be of worth to you. Of course, on the other hand, if the results on the kit are positive, the study has the added value of bringing attention to the kit. The study is focused to provide information for potential users of the kit. The basic methodology of the study is evaluative and focused on description and decision making. The two decisions the study addresses include:

a) Should work on the kit be continued?
b) Should the kit be mailed out (as is)?

Descriptive information provided by the study includes:
a) A demonstration of how the Evaluation of Instruction kit is used.

b) A demonstration of interim validation procedures."

"I must confess... I'm not sure I know what evaluative research is. I know that research is generally judged by the degree to which it demonstrates reliability, objectivity, and validity in the ways described by Campbell and Stanley. I expect good research to have an emphasis on generalizability of results. The purpose of research should be for testing old knowledge or creating new knowledge by supporting or refuting portions of a theory. Isn't evaluation just judging the worth of something?"

"I'm glad you're asking this. If you had read this study expecting research you are familiar with, you may have judged it with inappropriately weighted criteria."

"What are the standards for good evaluative research?"

"Reliability, validity, and objectivity are concerns in evaluative research, but other criteria are sometimes more important." I pick up a manual entitled "Similarities and Differences in the Research and Evaluation Processes." I know the man wants to believe me, but he would believe the words of a leader in evaluation more easily. I make the concession and read to him:
"William Michaels states the criteria for judging good evaluative research on page 32:

Fulfilling criteria of comprehensiveness and relevance to the decision-making task, affording efficiency and credibility, and maintaining to the extent possible standards of reliability and validity, and objectivity within a somewhat restricted context of endeavor moderated by important value-laden criteria in the educational-societal setting."

The specific criteria most important to a study in evaluation are linked to the purposes of the study and the values of the audience for the results. In this study the main audience is the potential field testers of the kit. A second audience is evaluators. A third audience is me—I have to make some hard decisions based partly on the results of this study. One of the most important audiences for the study is my doctoral committee. The study is mainly addressed to the potential field testers of the kit. Utility is therefore an important criterion for judging the kit, the workshop, and the entire study. Generalizability of results is not yet an important concern—generalizability of results is the appropriate concern at the field test stage rather than

in the early development stages. You expect your generalizability to change, increase, as the product nears completion and improves. Exploring generalizability is most relevant for the finished product.

5. Summary

I thought back over our talk: How I explained my interest in "education" and within that "evaluation" and within that a concentration on "higher education and continuing education" and within that "a kit" that is now under development. I then remembered explaining the importance of trying to assess the kit's likely usefulness before years of effort are devoted to it.

I looked over to the corner of the room. The man was doing push-ups. He slowly walked over to collapse in a chair. He picked up the dissertation and thumbed through it quickly to get an overview of what I did.
What I Did

1. Introduction and Review of Literature

"Tell me briefly what the dissertation contains," the man said. He put the dissertation down, leaned back, and relaxed in his overstuffed chair.

"As you have probably already noticed, most of the first pages are an introduction to the components of the study. The next few pages, the beginning of the review of literature, are an attempt to justify doing the study.

The rest of the review is almost an annotated bibliography. One reason for the format was to allow part of the dissertation to be a companion to the materials that are the focus of the study. The review was intended to provide potential users of the Evaluation of Instruction kit with alternatives to the kit and supplements to the kit. Since the literature relevant to the study is so vast, the format also allowed synthesis of a large quantity of material."

The man glanced over at the dissertation. "If you expect anyone to read your dissertation, why did you make it so long?"

"It's not really that long," I said, "over a third of it is Appendices. The Appendices are not supposed
to be read like a book. They are only reference material."

"Go ahead." This time the man let his head drop, he shut his eyes, and put his hands in his lap. I got a little worried. The dissertation is no novel. Maybe he just wanted me there to talk him to sleep.

"Don't worry," he said. "I really am interested. In fact, if the results of your study are worthwhile, I would like to try your kit to evaluate one of my classes."

Slightly reassured, I continued. "The Procedures section of the dissertation describes the study. It is also a justification of the methods and instruments used. Maybe you would like me to summarize the Procedures section."

A nod from the chair.

2. Procedures

"I found a group of willing participants and divided them by availability into two groups: one group received a kit through the mail and the other group was taught how to use the kit in a workshop. I wanted to find out if workshop training was completely necessary.

Before I could determine the importance of the workshop, I needed to know how well the workshop was implemented. For
instance, did the workshop participants really learn anything that would help them use the kit? I assessed the participants before and after the workshop to find out.

"But comparing the two modes of disseminating the kit wasn't the main purpose of your study was it?"

"No, the main purpose was to see if the kit was seen as useful and got used. I asked the participants (via a form) how useful they thought the kit was going to be for them. I also asked experts (via a form) how useful they projected the kit would be for potential users."

"But that doesn't tell you how much they actually would use the kit."

"That's where Ms. Armstrong comes in."

3. Enter Ms. Armstrong

"Much of what I have told you is typical of research studies. The next procedure, however, was a follow-up on both groups. Although a researcher collects follow-up information and interprets results, my doing that was inappropriate."

"You mean it was inappropriate for you to collect, analyze, and interpret the results on use of the kit?"

"That is exactly what I am saying. In addition, it was inappropriate for me alone to interpret the results on the impact of the workshop."
The man leaned forward in his chair; he was not going to accept this passively. "This is your dissertation, isn't it? How can it be inappropriate for you to interpret the results of your study?"

"I developed both the kit and the workshop. If I also judge the worth of both, I have a role conflict: as developer I need to believe that the kit and the workshop have a great deal of worth; as a judge I need to be impartial.

The solution was obvious. I collected information to monitor and revise the workshop; I also collected information for judging the impact of the workshop and the kit. Next, however, I hired an evaluator (Ms. Armstrong) to perform the follow-up study. She was also asked to interpret the information I had collected on the impact of the workshop and the kit. In other words, I separated my role as a developer from the role of impact evaluator.*

I gave the impact evaluator my data and my goals for the study, and allowed her freedom to assess the impact of the workshop and the kit. Two to four weeks after the workshop she contacted eight of the workshop participants and six of those receiving the kit in the mail."

*However, I also performed analysis and presented results and interpretations in order to display the outcomes of objectives #5 and #8 of this study clearly and cohesively.
"In other words, you conceptualized the impact evaluator tasks as a starting point for the impact evaluator to use as a frame of reference in beginning her impact evaluator role?"

"Right. In fact, it seems to me that whenever an educational evaluator is implementing a validation study (s)he should be careful in separating the developer's role from the impact evaluator's role. If the evaluator is developing the instruction that will be judged for worth, (s)he should consider taking the approach used in this study."

"You are exploring in quite a few directions. Let me see if I've got this straight:

You mainly wanted to see if the kit was valid for potential users; that is, you wanted to see if they would actually use the kit. So you gave them the kit and had someone see if (and how) they used it. Right?"

"Right."

"But you also wanted to see if the kit was valid in the eyes of experts. You sent the kit to experts so they could rate the contents."

"Right."

"Next in importance you wanted to see if you could just send the kit out—no training involved. Right?"
"Right."

"You wanted to demonstrate how the Evaluation of Instruction kit can be used. You wanted to demonstrate how to perform an interim validation study. You wanted to find out how successful the workshop was."

"Right."

"Hmm. I figure you had something more in mind on that workshop."

"Like what?"

"Like providing an example of what you feel to be good workshop evaluation; like showing how your kit can be used in actual practice; like watching close-up how others use the kit; like collecting information to improve your workshop; like showing you have a successful workshop that can go with the kit...."

"Okay, yeah, I had a lot tied into that workshop."

"It seems to me you were trying to collect a lot of information at the same time. What kind of results did you get?"

"Remember, I was the developer of the instruction. I did some analysis and interpretation on the implementation of the workshop, but you need to look at the Impact Evaluation Report by the impact evaluator to find out how useful the kit was. If you are not interested in the
technical side of the dissertation, you could even skip over to my Interpretations section on pages 165 through 177."

"Do you mean that the rest of this dissertation, except for the Interpretations section, is a technical report?"

"Yes."

"Reading a technical report is not my idea of an enjoyable evening. You collected information for all kinds of purposes--how am I supposed to know what to look for?"

4. Objectives

"You could use my objectives to guide you: Objective number one. The Review of Literature chapter constitutes a descriptive list of resources for evaluation of instruction and is the outcome of objective number one.

Objective number two. The kit accompanying the dissertation constitutes the outcome of objective number two; the kit is described on pages 34-35.

Objective number three. The workshop described on pages 36-38 constitutes the outcome of objective number three.

Objective number four. Evaluating the impact of
the workshop was the most complicated part. I was looking for congruence reliability—I wanted to see if I got the same results from many kinds of indicators. I wanted a wholistic idea of what kind of impact the workshop had. Any one set of results could be deceptive."

"Why don't you just give me a list of things to look for?"

"I would consider the workshop a success to the degree that I found:

a. Significant gains by participants using pre-post measures on their skill, knowledge, attitude toward evaluation, and achievement (self-rated).

b. Better evaluation instruments by participants after the workshop than before the workshop.

c. Participants seeking out evaluation specialists for help.

d. Many participants using the kit after going through the workshop.

e. More participants with an overall evaluation design after the workshop than before the workshop.

f. Positive judgements toward the workshop from participants."

"You also want me to notice that the study is more
than just finding results on a kit, don’t you?"

"Yes. Objective number five. I would like for you to notice that all of the instruments I used to assess the workshop were based on instruments in the kit. Even the framework used in collecting data was taken from the kit. Objective five is to demonstrate the utility of the Evaluation of Instruction kit through its application to a workshop. The demonstration includes more than just assessing the impact of the workshop; it also includes collecting diagnostic information.

Objective number six. I would like to assess use of the Evaluation of Instruction kit through its application in evaluating the Evaluation of Instruction workshop.

Objective number seven. I would like to assess the interim validity of the Evaluation of Instruction kit.

Objective number eight. The overall design of the study was meant to be of help in performing interim validation studies on other educational products. Objective number eight is to demonstrate a potentially transferable framework for interim validation of an educational product. The following constitute components of the interim validation: A successful application of the product (the kit), the perceived utility of the product, the projected utility of the product as a whole and in parts, assessment of the
training in use of the product, assessment of use of the product without training, and assessment of use of the product after training."

"I think I've heard enough. Maybe I should just read for awhile," the man said.

"The next chapter is a review of literature. The review includes a descriptive list of resources for use with evaluation of instruction; the review is also a supplement to the Evaluation of Instruction kit. Chapter II, the review, is the outcome of the first objective of the study. I make frequent references to the Evaluation of Instruction kit in the review. Let me give you a brief overview of the kit before you read further. Let me also give you a brief overview of the workshop that was conducted to train potential users of the kit. The review of literature should make more sense after you have read the two overviews.

Before you get started - who the heck are you?" I was too late; he was already deeply absorbed in my overview of the Evaluation of Instruction kit.
Overview of the Evaluation of Instruction Kit

The Evaluation of Instruction kit by the writer is composed of three components:

a) A manual entitled "How to Use Evaluation of Instruction Kit" was designed to orient the user to the kit, provide a prototype set of evaluation instruments and an evaluation design, and provide a framework for using the entire Evaluation of Instruction Kit.

b) A manual entitled "Evaluation of Instruction" was designed to provide evaluation instruments and guidelines in compact form; and

c) The Evaluation of Instruction Kit consists of the above two manuals, a copy of each page in "b" above, a folder for each page so that retrieval of information will be easier, and a plastic letter file to contain the above.

The framework for the evaluation kit was constructed as a synthesis of the more popular program evaluation frameworks, including the CIPP model, the CSE model, and the Countenance model.*

The instruments and guidelines within the kit were designed to serve as samples to fit each category.

*See Review of Literature.
of the conceptual framework. The developer of the kit attempted to select or develop instruments that were easy to use (short), easy to analyze, easy to interpret, and broad in scope. Selections were made on the basis of a review of literature and experience with evaluation instruments for over one hundred workshop presentations by others.
Overview of the Evaluation of Instruction Workshop

The Evaluation of Instruction workshop was conducted on June 11, 12, and 13, 1973. It was created to provide training for potential users of the Evaluation of Instruction Kit. Workshop participants were instructors, developers, and administrators who were involved in providing training designed to help individuals with developmental disabilities.

The participants in the workshop represented a broader range of backgrounds and needs than intended by the workshop developer; however, the participants were divided into homogenous work groups to accommodate differing needs.

The developer of the workshop attempted to change the workshop as a result of evaluation feedback obtained during the workshop. The evaluation specialists, Ms. Unhai Ahn and Mr. Gary Milczarek, assisted the developer of the workshop in leading work groups. The workshop generally followed the agenda,* but actual components lasted longer than intended. The second and third days of the workshop were shortened by an hour each at the prior request of the participants. (They wanted to sleep longer in the morning.) Two presentations were added to the agenda as a result of expressed participant concerns: one presentation was on evaluation jargon and

*See pages 203-204.
the second presentation was on evaluation models.

Objectives for the workshop included the following:

a) develop an instrument for analysis of instructional materials

b) develop an instrument to perform a needs assessment

c) develop a criterion referenced instrument

d) develop an instrument to monitor the instructional process

e) develop an instrument to collect student reactions

f) develop an instrument to assess the impact of instruction, including unintended outcomes

g) develop an instrument to assess the long-term effects of instruction

h) develop an evaluation design or data collection framework

i) locate resources for data analysis

j) validate an instrument

k) interpret evaluation results

l) report evaluation results

Activities of the workshop were focused on instrument construction, critiques, data analysis, and reporting by participants. The workshop was intended to give participants a chance to develop an awareness of the kit, an interest in the kit, comprehension of the kit, a trial of the kit, and a chance to evaluate the kit. The last stage in use of the kit, adoption of the
kit, was left up to the participants after the workshop. Support in using the kit after the workshop was offered in the form of free consulting service by the workshop presenter.
Chapter II
REVIEW OF LITERATURE

The review formed a foundation for development of a kit for evaluation of instruction. After the development of the Evaluation of Instruction kit, the review still serves to justify the development and assessment of the kit; it also offers alternatives to the kit.

The review addresses two questions:

A. Is the Evaluation of Instruction kit likely to be a useful product for the field of educational evaluation; is the kit important enough to justify a conscientious assessment during its early development?

B. What alternatives to the kit are available?
Need for Kit and Validation

The review addresses question A first: Is the Evaluation of Instruction kit likely to be a useful product for the field of educational evaluation; is the kit important enough to justify an early assessment?

Support for the importance of assessing the kit includes the following opinions and findings:

Worthen and Sanders (1973) claimed:

Perhaps the greatest impediment to progress in evaluation is lack of unequivocal knowledge about many of its critical components. For example, there is little or no data-based information about the relative efficiency of alternative evaluation plans or techniques.¹

The kit contains techniques and a framework for building an evaluation plan. If Worthen and Sanders are correct, then collecting data on the kit is an important activity.

The Phi Delta Kappa National Study Committee on Evaluation (Stufflebeam, et. al., 1971) described the most important needs in educational evaluation. The

areas needing greatest attention were:

a) appropriate instruments and designs,
b) mechanisms for organizing, processing and reporting evaluative information, and
c) specification of the kind of evaluative information which would be most useful.\(^2\)

The kit addresses each of the above needs.

Cronbach (1973) quoted Bloom (1961) as follows:

> When teachers have actively participated in . . . selecting or constructing evaluation instruments, they return to the learning problems with great vigor and remarkable creativity.\(^3\)

The kit is an attempt to help instructors with selection or construction of evaluation instruments.

Cronbach (1973) also stated that "the benefit (of the above) is attributed to thinking about data to collect."\(^4\)

The kit provides a wide range of alternatives in data collection.

Bensberg (1969) revealed:

> Evaluation of our training programs and its effectiveness is difficult and requires skills which many of us as instructors do not have. Also, it requires


\(^3\)Benjamin Bloom, "Quality Control in Education," Tommorrow's Teaching (Oklahoma City: Frontiers of Science Foundation, 1961).

staff time which is frequently not available. However, more and more we are looking at this aspect of training and attempting to build in more evaluation methods as an essential part of training.  

The kit was designed to decrease the difficulty in building an evaluation plan.

Cronbach (1973) stated:

Evaluation used to improve the course while it is still fluid, contributes more to improvement of education than evaluation used to appraise a product already placed on the market.  

The emphasis of the kit is on improvement of instruction.

Rossi (1972) stated:

Evaluation has been popular in terms of federal legislation, but it has been less than a great success in its outputs.

The kit was designed to improve evaluation outputs.

Bergman (1972) conducted a review of literature (1962-1972) of evaluation of instructional systems. One of his

---


6Op Cit.

conclusions was that "systematic approaches to evaluation and course development are receiving more and more attention."\(^8\)
The increased attention suggests that a need exists. The kit is a systematic approach to evaluation and course improvement.

Musella (1966) found in a review of literature that over 3,000 studies have been conducted in an attempt to identify the determinants of effective teaching; however, he found that problems related to systematic evaluation of instruction have received much less attention. The statement implies that a framework to synthesize findings has been lacking or inadequate.

The kit is an attempt to provide a framework for systematic evaluation of instruction.

Astin and Lee (1966) reported on a study by the American Council on Education. The Council conducted a survey on evaluation of teachers in all higher education institutions. The Council received 1,110 useable completed questionnaires. Evaluation of teacher effectiveness was most frequently performed by the dean and department chairman.

Since deans and department chairmen are not necessarily the best judges of teacher effectiveness, additional sources of information are needed. The kit was designed to fill the need for alternatives.

---

Walters (1973) conducted a study on the use of student evaluation of instruction at The Ohio State University by questioning department chairmen. He concluded:

In summary, there are strong indications that a highly flexible system [of student evaluation of instruction] . . . would be appreciated by many [instructors].

The kit was designed to provide instructors with a flexible system for evaluation.

Smock and Crooks (1973) emphasized that "... the need grows for more workable systems directed toward aiding faculty as they seek instructional change."10

The kit was designed to assist faculty with instructional change.

The report of the Centennial Commission on the Future of The Ohio State University (1972) included the statement that "greater systematization must be provided, especially for appraising nonresearch activities [of faculty]."11

9Garrison E. Walters, "Student Evaluation of Teaching at the Ohio State University," (Unpublished draft, Ohio State University, 1973), p. 5.


The kit is an attempt to systematize the appraisal of the major nonresearch activity of faculty.

Three hundred and twenty-seven Ohio citizens and thirty-four alumni answered the question: "What is the most important goal you see for higher education to achieve in Ohio?" The most frequent response was, "Improve the quality of education."\(^{12}\)

The kit was designed to facilitate improvement of instruction.

**Summary**

In summary, the kit was designed to meet important stated or implied needs in evaluation of instruction. The kit was designed to:

a. Provide appropriate evaluation instruments and plans.

b. Provide mechanisms for organizing, processing, and reporting evaluative information.

c. Specify the kinds of information which might be useful.

d. Facilitate selection or construction of evaluation instruments by instructors.

e. Encourage instructors to concentrate more on data collection.

f. Decrease the difficulty of building an evaluation plan.

g. Improve the utility and face validity of evaluation results for instructors.

h. Encourage improvement of instruction as instruction progresses.

\(^{12}\)Ibid.
Alternatives to Evaluation Kit

The second question that the review addresses is: What alternatives to the kit are available?

An answer to the above question covers at least the following types of efforts in educational evaluation:

a. Alternative evaluation frameworks

b. Complementary guides
   1. Training packages
   2. Guides for an overview of evaluation
   3. Guides for performing evaluation

c. Alternative evaluation instruments
   1. Compendia of listings and reviews
   2. Compendia of evaluation instruments

d. Complementary instrument construction guides
   1. Guides to constructing objectives
   2. Guides for evaluation of students
   3. Guides for student evaluation of instruction

e. Review of Alternatives Summation
Alternative Evaluation Frameworks*

Sanders and Cunningham Framework

Sanders and Cunningham (1973) provided one of the most recent frameworks for evaluation of instruction. They provided a matrix with evaluation activity and sources of information. Procedures and techniques fit inside the cells of the matrix:

<table>
<thead>
<tr>
<th>Sources of Information</th>
<th>Pre-Developmental</th>
<th>Evaluation of Objectives</th>
<th>Interim</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>External</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 1**

Sanders and Cunningham asserted that use of the above formative framework "should alleviate many anxieties of developers about potential surprise criticisms in a summative evaluation."\(^{13}\) If an instructor could wade through the jargon, he might find the framework quite useful.

Comparison With Kit**

The Pre-developmental activities and the Evaluation of Objectives activities would logically occur before instruction began; therefore, they would probably

---

* Alternatives to the Evaluation of Instruction Kit

** Evaluation of Instruction Kit

fit into the Before the Session category of the Evaluation of Instruction Kit. Interim activities might occur when the instruction was in progress, so the category is similar to the During the Session category of the kit. Product evaluation activities would logically occur after the instruction had occurred, so the category might be seen as similar to After the Session.

The kit has no categories analogous to the Sources of Information categories of Sanders and Cunningham. The kit has, however, instruments and guides as categories that would fit into the cells provided by the Sanders and Cunningham matrix.

The Sources of Information categories add an element of sophistication and utility to the Sanders and Cunningham framework the kit's framework lacks; however, the kit was not designed for an audience sophisticated in evaluation concepts or techniques. The third dimension might be seen as too much of a refinement for the intended audience.
Stake Framework

Stake (1967) presented one of the most popular data collection frameworks:

<table>
<thead>
<tr>
<th>Rationale</th>
<th>Intents</th>
<th>Observations</th>
<th>Standards</th>
<th>Judgements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Antecedents</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 2**

Gish (1971) used the Stake framework to generate a data collection matrix for evaluation of an early childhood program.
STAKE FRAMEWORK
(modified by Gish)

<table>
<thead>
<tr>
<th></th>
<th>Input (Antecedent)</th>
<th>Process (Transaction)</th>
<th>Output (Outcome)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td>attitude and goals</td>
<td>activities related to program</td>
<td>satisfaction</td>
</tr>
<tr>
<td>Teacher</td>
<td>objectives</td>
<td>classroom milieu and teacher</td>
<td>satisfaction</td>
</tr>
<tr>
<td>Child</td>
<td>entry behavior</td>
<td>child class interact</td>
<td>change</td>
</tr>
</tbody>
</table>

FIGURE 3
Comparison With Kit

Since the kit deals only with adult education, collection of information on parents would not be a high priority. All other indicators in Gish's matrix are contained in the kit -- except one. The kit does not contain an indicator of teacher satisfaction as a product or outcome of the instruction. An instructor might collect useful information for himself by filling out one or several of the instruments in the kit using his own satisfaction as the criterion.

The Antecedents category is equal to the Before the Session and Beginning the Session of the kit; the Transaction category is equal to Throughout the Session; and Outcomes are equal to End of the Session and After the Session in the kit. Intents, Observations, Standards, and Judgements, are not, however, collected systematically in all categories of the kit. Rationale is not dealt with in the kit.

Hammond Framework

The framework presented for program evaluation by Hammond (1973) is three dimensional:
FIGURE 4

HAMMOND FRAMEWORK

BEHAVIOR

Psychomotor Domain
Affective Domain
Cognitive Domain

INSTRUCTION

Organization
Content
Method
Facilities
Cost

INSTITUTION

Student
Teacher
Administration
Educational Specialist
Family
Community
Relevant elements of the Instructional Dimension
of Hammond's evaluation framework include:

1. Organization
2. Time
3. Space

Content (Topics)
A. Methodology
   1. Teaching Activities
      a. Lecture
      b. Discussion
      c. Question-Answer
      d. Committee
      e. Round Table
      f. Symposium
      g. Drill
      h. Homework
      i. Review
      j. Individual Supervised Study
      k. Resource Person(s)
      l. Field Trips
      m. Inquiry
      n. Debate
      o. Media
2. Types of Interaction
   a. Teacher-Student
   b. Student-Student
   c. Media-Student
   d. Teacher-Teacher

3. Learning Theory

B. Facilities
   1. Space
   2. Special equipment
   3. Expendables

C. Cost (Money for facilities, maintenance, and personnel)

Comparison With Kit

The structure produced by the Behavioral, Instructional, and Institutional variables of Hammond is much more comprehensive than the structure presented in the Evaluation of Instruction Kit. The kit not only uses a much less comprehensive set of variables in its framework for evaluation of instruction; but it also gives much more focus to areas not obviously touched upon by Hammond. For example, outputs of instruction are not clearly presented in Hammond's framework.
McGuire Framework

McGuire (1967) developed a data collection plan:

**SOURCE OF INFORMATION**

<table>
<thead>
<tr>
<th>Expert Opinion</th>
<th>Participant Reaction (Process Analysis)</th>
<th>Performance (Outcomes)</th>
<th>Requirements (Inputs)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teachers: Background</td>
<td>Satisfactions:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Students: Background,</td>
<td>Teachers, Community,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ability, interest</td>
<td>etc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Setting: climate for</td>
<td>Cognitive Achievement:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>learning in all do-</td>
<td>Immediate and Long-Run</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mains</td>
<td>Affective Changes:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Materials: texts, etc.</td>
<td>Immediate and Long-Run</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Methods: Teacher-</td>
<td>Psychomotor</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Student</td>
<td>Achievement:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Transactions</td>
<td>Immediate and Long-Run</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Costs: Time, money</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Goals Sacrificed</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 5**

---

McGuire stated the purposes of her framework: 15

1. Document entry behaviors.
2. Specify setting.
3. Describe transactions and materials (the treatment).
4. Summarize the product.
5. Judge the merit of the product.

Comparison With Kit

McGuire denoted four types of inputs not described in the Evaluation of Instruction kit: 16

1. Background of teacher
2. Setting or climate for learning in all domains
3. Costs in time and money
4. Goals sacrificed

Since the Evaluation of Instruction kit is intended for use by instructors, a description of a teacher's background might not be useful. However, the instructor might gain useful information by filling out the same forms as his students and comparing his criteria, preferences, and expectations with theirs. Costs in time and money and Goals Sacrificed could have been items on the Analysis of Materials instrument in the kit but were seen as less high priority for instructors than the other items chosen.

15 Ibid., p. 94.
16 Ibid.
Sheier (1973) developed a manual containing an annotated taxonomy of modes for collecting evaluation information. The modes include:

A. Expert Opinion
   1. Consulting
   2. Position Papers
   3. Meetings

B. Studies
   1. Tertiary Sources
   2. Secondary Sources
   3. Primary Sources
   4. Computer Simulation

C. Surveys
   1. Public Opinion
   2. Questionnaires
   3. Interviews
   4. Unobtrusive Measures
   5. Nonstandardized Tests

D. Behavior Observation
   1. Subjective Observation
   2. Controlled Observation (Judges)
   3. Frequency Counting
E. Standardized Tests
   1. Personality
   2. Attitudes and Values
   3. Intelligence
   4. Achievement
   5. Skills
   6. Physical Health Related

F. Laboratory Measurement
   1. Environment Simulation
   2. Control Group

Comparison With Kit
The kit contains a small sample from the Sheier taxonomy. The taxonomy might be helpful as a stimulus for instructors looking for new modes of data collection.

Stufflebeam Framework
Stufflebeam's CIPP model was one of the more popular frameworks presented by Worthen and Sanders. Stufflebeam et al. (1971) attempted to draw the major conceptions of evaluation together into a single book which not only gave a unifying framework for evaluation, but also outlined guides for performing evaluation. The major components of the Stufflebeam model are:

   context evaluation (C)
   input evaluation (I)
   process evaluation (P)
   product evaluation (P)
Wolf and Smith (1973), Milczarek (1973), and Sanders et al. (1972) have applied the CIPP framework to evaluation of instruction in higher education.

Wolf and Smith described a model of curriculum evaluation based on the CIPP model of program evaluation, but they were not very specific.

Milczarek purported that he based much of his evaluation design for the EPEC instructional package on CIPP. His evaluation system requires obtaining use of an observation schedule, use of a formative evaluation instrument throughout sessions, a mastery test, a session questionnaire, and a final questionnaire.

The unit of measurement for the Milczarek design is "learning activities." Each learning activity is assessed in terms of content, objectives, environmental conditions, media, materials, feedback, time and placement.

Criteria for each of the learning activities are stated; the decision rule, data source, and instrument used to assess each learning activity, are also specified.

Milczarek's design is very meticulous. Although it can be recommended highly for formative evaluators, the design requires skills a typical instructor does not have. The design was built for specialists in
evaluation, not for instructors.

The Center Instructional Development (CID) model by Sanders et. al. delineates activities, major decisions, and feedback loops for each of the CIPP categories. Most instructors should be able to read and understand the flow chart of activities, but the model was designed mostly for development specialists. The delineation of tasks associated with the model is very comprehensive and should prove of use as a stimulus for instructors.

Comparison With Kit

Context evaluation logically would preceed instruction and might be seen as somewhat analogous to Before the Session and Beginning the Session categories. Input evaluation has no exact counterpart in the kit, but would probably fit with Beginning the Session. Process evaluation would fit with Throughout the Session. Product evaluation would fit with End of the Session and After the Session.

Milczarek's model centers much on a data collection framework; indicators Milczarek used could be seen as supplementary to or examples of indicators suggested in the kit.

The Sanders CID model might be very useful as a complementary resource to the kit.
Alkin Framework

The Alkin (1972) framework for program evaluation is similar to the Stufflebeam conceptualization. In fact, Alkin states that "Stufflebeam's work contributed substantially to our thinking." Alkin described context evaluation as similar to his system assessment and program planning stages. Alkin's program implementation and program improvement stages were seen as related to Stufflebeam's input evaluation, as well as Stufflebeam's process evaluation. Program certification was seen as similar to Stufflebeam's product evaluation. Alkin's framework, then, includes the following components:

Systems Assessment
Program Planning
Program Implementation
Program Improvement
Program Certification

Like Stufflebeam, Alkin's model was portrayed in a flow chart format with decision points between categories.

Comparison With Kit

Alkin's categories of Systems Assessment and Program Planning seem similar to Before the Session and Beginning of Session; however, Beginning of the Session is considered part of the "treatment" in the kit framework and Alkin's categories refer to activities taking place before a program is implemented. Program Improvement is similar to During the Session. Program Certification is similar to After the Session.

Provus (1971) described four stages of development of a program:

1) design,
2) installation,
3) process,
4) product.

Provus also developed a taxonomy of components of program content. The taxonomy emphasizes transactions, time estimates, and cost estimates. The taxonomy links Acquisition time and Fixed Costs with:

Input:
Staff qualifications by position
Staff preprogram training
Student selection criteria
Student entry behavior

Media

Facilities

Administrative conditions

Operating Time, Cumulate and Storage Time and Operating Costs are linked with:

Process:

Standard transactions with:

Students

Staff

Media

Facilities

Staff transactions with:

Staff

Students

Media

Facilities

Administration

Others

Student-staff transactions relative to objectives.

Operating Time, Cumulate and Storage Time, and Cumulate and Storage Costs are linked with:

Outputs:

Enabling Objectives (EO)

Terminal Objectives (TO)
Ultimate Objectives (UO)

Interrelationship between EO's, TO's, and UO's.

Comparison With Kit

Since the kit deals with evaluation of instruction rather than evaluation of instructional programs, monitoring time and cost did not seem high priority. A major alternative to the framework of the kit is presented in the Outputs of the Provus taxonomy. Outputs are measured in terms of objectives alone in the Provus approach, but the objectives are divided into types so that measurement might be a rather comprehensive and ambitious undertaking.

The Inputs category of Provus is somewhat similar to the Before the Session and Beginning the Session categories in the kit. The Process category is similar to the Throughout the Session category in the kit. The Outputs category is similar to the End of Session and After the Session categories of the kit.

Synergistic Framework

Hunter and Schooley presented a program evaluation model relevant to evaluation of instruction. The model contains the following components:
Policy Domain
   a. Establishment of system goals
   b. Development of performance objectives

Program Development Domain
   a. Needs Assessment
   b. Instructional System Analysis
   c. Program Modification

Instruction Domain
   a. Program implementation
   b. Formative evaluation

Feedback Domain
   a. Summative evaluation
   b. Recycle within the model

Comparison With Kit

The kit contains all of the elements of the synergistic model.

Smock and Crooks Framework

Smock and Crooks (1973) developed a framework for evaluation of instruction in higher education. Smock and Crooks basically advocated collecting three levels of evaluation information:

Level I: Summative, interdepartmental information
Level II: Summative intradepartmental information

Level III: Formative, diagnostic information

Comparison With Kit

The kit focuses mostly on what Smock and Crooks would call Level III evaluation. The kit does not address the political considerations in dissemination of evaluative information but rather assumes that both formative and summative information will be helpful to the instructor in Level III evaluation.

Kunkel-McElhinney Framework

Rush and Fifer (1973) described the Kunkel-McElhinney model of curriculum evaluation. The model was presented in very general terms but seemed to mainly include a statement of intention and three modes of collecting information:

Every effort is made to convey to the population being assessed that the evaluation team is coming to lend assistance in conducting a self-study.  

Data are collected from

1) written reports
2) structured interviews
3) questionnaires

---

Comparison With Kit

The framework of the kit seems to encompass the components of the Kunkel-McElhinney model.

Instructional Development Frameworks

Cronbach (1963) has claimed that evaluation and instructional development are closely interwoven. Models for instructional development seem to generally include evaluation as an integral part of the model.

The development model used by the Southwest Educational Development Laboratory (1970)* includes evaluation throughout the development process:

1. Context analysis
2. Conceptual design
3. Product design
4. Pilot test
5. Field test
6. Marketing and Diffusion

* The Southwest Educational Development Laboratory. The Calipurs: Planning the System Approach to Field Testing Educational Products. (Austin, Texas: Southwest Educational Development Corporation, 1969)
The lesson function chart presented by Smith (1966) can be seen as a matrix with objectives across the top and tasks down the side; symbols are entered into the cells according to a legend. The chart can be viewed as an evaluation data collection framework.\(^{19}\)

**LESSON FUNCTION CHART**

<table>
<thead>
<tr>
<th>Objectives</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 6**

Smith also included "A Check List for Evaluating Training" in his Appendix.\(^{20}\)

Banathy's (1968) chart "The Design of Instructional Systems"\(^{21}\) includes evaluative information throughout the system.

**Comparison With Kit**

The above mentioned are only examples of instructional development frameworks that could double as evaluation frameworks. They differ from the kit in that they

---


\(^{20}\)Ibid., pp. 83-5.

generally specify developmental activities as well as evaluative activities. The kit assumes that development is already somewhat complete and does not address developmental activities with much strength.
Complementary Evaluation Guides

1. Training Packages

Several training packages and many books complement the purpose of the Evaluation of Instruction kit.

Training Packages:

The Teaching Improvement Kit (Popham) was designed to facilitate teaching competency. Evaluation of teaching performance was based on

- post tests of student performance and
- rating on student interest.

The Prentice-Hall Teacher Competency Development System (Popham and Baker) contains twenty-seven units, including twenty-six booklet modules and an Administrators' Guidebook. The kit was designed to train teachers in areas that include assessment.

The Higher Education Measurement and Evaluation Kit (Pace, 1971) was developed to facilitate the study of consequences of higher education.

Evaluation Workshop I: An Orientation (Klein, 1971) was designed to help the user gain a wide range of evaluation skills.

Fehr-Practicum (Leverne S. Collet) is now under development. It is a computer based simulation designed to teach research and evaluation skills.
Qualitative Educational Evaluation (Scriven) is under development to assist the user in performing evaluation activities.

An Introduction to the Theories, Functions, and Methods of Education Evaluation (Stufflebeam) is under development to give the user a broad overview of the field of educational evaluation.

A Training Program in Educational Development, Dissemination, and Evaluation (DD and E) (Banathy) is under development for individuals in or going into DD and E.

Series Five: Evaluation will address the following topic areas:

a. Module 5.1: Introduction to Evaluation
b. Module 5.2: Tests and Measures in DD and E
c. Module 5.3: Design of Instruments
d. Module 5.4: Development and Implementation of Evaluation Plans
e. Module 5.5: Evaluation Problems

Evaluation in Education (Gottman and Clasen, 1972) was designed to train the user in performing the following tasks:

a. Do a needs assessment
b. Write measurable objectives and design measurement procedures
c. Generate a flow chart
d. Design and use quality control procedures

A Technology for Developing Instructional Materials (Gropper, 1973) covers "task analysis, stating objectives, developing tests, planning simulations, formulating strategies, and development, trial, and revision of materials." \(^{22}\)

Preparing Instructional Objectives (Mager, 1962) assists the reader in understanding how to write good instructional objectives.

Research Utilization and Problem Solving (Jung, Pino, and Emory, 1970) contains instruments and strategies relevant to needs assessment and assessment of small group processes.

Evaluation Workshop: Needs Assessment (Klein) is in the field test stage.

A Seminar and Training Program in Needs Assessment and Goal Development (Morgan, 1972) was developed on two levels: one level to help laymen with the topic area and the other level to assist the evaluation specialist.


Item Writing: A Programmed Text of Rules for Writing Objective-Type Test Items (Rupiper, 1964) -- the title is self-explanatory.

Hills (1973) is developing an instructional system in measurement for teachers. The topics are:

a. Grading and marking
b. Building classroom tests
c. Choosing, administering, and interpreting standardized tests
d. Using tests to improve instruction

Ahman and Glock (1971) produced a useful booklet on contracting test items for assessment of achievement.

Introduction to Measurement: Instructional System #003 (Pegram, 1972) was designed to address statistical problems in measurement.

Introductory Statistics for the Behavioral Sciences (Young, 1965) is a partially programmed text on elementary statistics.

An Interviewing Training Model (Haller, 1973) was designed to develop interviewing skills.

Contracting for Evaluation Services (Klein, 1971) takes the user through exercises to develop his skill in writing a contract for evaluation services.

Evan Baker's materials should assist the instructor in conceptualizing a data collection plan during development of his instruction.
2. Guides for an Overview of Evaluation

Worthen and Sanders (1973) developed a book which, in their words, "includes both the most promising conceptual frameworks proposed for educational evaluations and practical considerations in conducting such evaluations."24

Of particular worth within the book is a chart25 describing and comparing major evaluation frameworks. Although the descriptions may not be entirely accurate, the matrix should afford the novice an overview of evaluation.

Worthen and Sanders also presented a very useful table of strengths and weaknesses of eighteen basic methods of collecting evaluation data.26

Stufflebeam, et al. (1971) generated a book to synthesize knowledge in the area of evaluation and generate a framework and guidelines for performing evaluation. The authors stated their objectives as follows:

This book . . . is organized to be responsive to four objectives . . .

---

26 Ibid., pp. 286-7.
Its purpose is to expose five problem areas . . . to identify and assess approaches to deal with these . . . to synthesize a new definition and methodology of evaluation resulting from the assessment . . . and to provide operational guidelines for implementing the proposed new approach.27

Tyler, et al. (1969) developed a book that the authors wanted judged "in terms of the impetus that it will give to a rethinking of the broader aims and purposes of evaluation, to the consideration of new and untried methods, and to the refinement of procedures that have been developed."28

Wittrock and Wiley (1970) edited the report of a symposium designed to bring "new approaches to old problems and issues of evaluating instruction."29

Weiss (1972) brought together a set of papers to assist the reader to "learn what evaluation is all about and how its results fit into the process of program decision-making."30


3. Guides for Performing Evaluation

Evaluating Instruction by Popham (1973) should be helpful to the instructor. An earlier work by Popham (1972) contains twenty rules for the educational evaluator to follow.

Faculty Handbook, Part II: Improving Teaching (Dept. of Ag.) includes Allport's teacher evaluation plan, a checklist for self-evaluation by the instructor, and an appendix of teaching techniques according to Pettman (1973).

The handbook by Wandt and Brown (1965) was designed for use by teachers and was focused on measurement problems and methods.

Smock and Crooks (1973b) cited Eble's (1970) book as having good suggestions in the areas of systematic selection, collection, and use of evaluative data related to instruction.

Smock and Crooks (1973b) cited Miller's (1972) book for its useful information related to evaluation of instruction.

The handbook by Mildebrand, Wilson, and Dienst (1971) contains a helpful outline of major decisions to be considered by those responsible for evaluation of instruction.

Bloom and Hastings (1971) provide much help in assessment of instructional outcomes.
The handbook by Isaac and Michael (1971) provides guidelines, methods, and techniques for analysis of data, formulating the problems, using research design, and generating instruments.

Sheier's (1969) manual covers the following areas: planning evaluation, stating objectives, research design, implementation, analyzing and interpreting results, reporting results, and "translating the findings into curriculum improvements." 31

The manual also contains a "taxonomy of procedure for selection of data collection instruments, a brief description of statistical tests and suggestions for questionnaire items." 32 Appendix C contains an outline of statistical tests for laymen.

Miles in Measuring Human Behavior, judged Miller's (1970) handbook as follows:

As a general research handbook for the social (primarily sociological) scientist, this volume is almost unparalleled; it is nearly a do-it-yourself kit for researchers. 33


32 Ibid.

Fear's (1973) book gives helpful hints for use with follow-up interviews and debriefing participants of workshops.

Preparing Evaluation Reports: A Guide for Authors (USGPO) was directed toward evaluators of public school programs.

Alternative Evaluation Instruments

The Evaluation of Instruction Kit contains only a miniscule of generalizable tests appropriate for evaluation of instruction. Many other tests may be an appropriate and highly useful part of an evaluation plan. Two types of sources may be useful to the instructor to supplement the instruments available in the kit. One type of source for instrument selection is compendia of listings and reviews of instruments; the other type of source for instrument selection is compendia of instruments. A few examples of each type are listed below:

1. Compendia of Listings and Reviews

Lake, Miles, and Earle's (1973) excellent text contains reviews of assessment instruments, assessment instrument compendia, and assessment instrument compendia reviews.
The Mental Measurement Yearbooks edited by Buros are classic reference sources. The Seventh Mental Measurement Yearbook (1971) contains reviews of 1,159 tests (a yearbook is not published every year; only seven have been published so far).

Buros (1961) has also edited a bibliography of 2,126 tests. The reference book contains a technical recommendations section and a publisher and distributor index.

2. Compendia of Evaluation Instruments

Published Standardized Tests: An Annotated List for Junior Colleges contains 330 tests derived from 23 test publishers. The tests in Scibell's (1967) book cover the content areas of ability, pupil skills, interest inventories, and achievement.

Shaw's (1967) Scales for the Measurement of Attitudes is a compendium of 176 attitude scales with critical reviews. Miles in Measuring Human Behavior considers Shaw's work "a model of good compendia of instruments."\(^{34}\)


\(^{34}\)Ibid., p. 368
Werdell's (1967) manual entitled Course and Teacher Evaluation is designed to help students initiate their own student evaluation of teaching. It includes many instruments for student evaluation of instruction.

MIRRORS FOR BEHAVIOR: AN ANTHOLOGY OF INSTRUMENTS is edited by Simon and Boyer (1968) and contains many instruments useful in higher education and in continuing education.

Culbertson's (1972) package of simulation materials contains many instruments for diagnosis of group simulation outcomes.

DIAGNOSING CLASSROOM LEARNING ENVIRONMENTS by Fox, Luszki, and Schmuck (1966) includes 23 diagnostic tools for teachers.
Complementary Instrument Construction Guides

Evaluation of instruction without objectives can be difficult. Many sources were designed to assist with the development of objectives.

However, most sources on evaluation focused on evaluation of students by instructors via formal tests or evaluation of instruction by the students via summative assessment forms. Systematic observation of the teaching process is another popular focus, but the focus seems mostly for research rather than evaluation purposes. No source was found that gave guidelines and procedures for the more general area of evaluation of instruction.

Instrument construction guides will therefore cover three areas: guides to constructing objectives, guides for evaluation of students, and guides for student evaluation of instruction.

1. Guides to Constructing Objectives

Mager's (1962) book is a very popular source for teaching how to write objectives.

Armstrong's (1968) handbook is similar to Mager's in style, content, and format.

Bloom's (1956) taxonomy was designed to be of help in construction of objectives for the cognitive domain.
Krathwohl, Bloom, and Masia's (1956) taxonomy was designed to be of help in the construction of objectives for the affective domain.

Some of Metfessel and Michael's (1967) criterion measures may be found useful when paired with objectives from the cognitive, affective, or psychomotor domains.

Pascal and Roid (1973) tested a method for developing university departmental objectives. Several groups were formed to generate objectives, consultants rewrote objectives, group members responded to the total list of objectives, consultants screened and rewrote objectives, and finally instructors rated each objective on relevance.

Popham (1972) stated procedural guidelines for developing objective-based tests. For instance, Popham listed and discussed criteria for selection of objectives for tests.\(^{35}\)

Similarly, Popham listed and discussed criteria for selecting learner behaviors which should guide selection of test items.

Appendix A of Popham's paper gives format guidelines and Appendix C supplies a summary of development procedures useful in construction of tests of students.

\(^{35}\)James W. Popham, Procedural Guidelines (Los Angeles, California, 1972), p. 5.
2. Guides for Evaluation of Students

Ebel (1951) provided a rather complete set of guidelines (advice) for writing good test items.

Ebel's (1972) chapter "How to Judge the Quality of a Classroom Test" might be of particular interest. For instance, Ebel's form for analyzing a test\textsuperscript{36} should be useful to the instructor. Ebel defined types of validity in terms of classroom tests rather than in general terms.\textsuperscript{37} Ebel also supplied criteria for assessing a good achievement test instrument.\textsuperscript{38}

*Constructing Evaluation Instruments* by Furst (1968) contains two "cookbook" chapters: Chapter Eight, "Constructing Items to Fit Specifications," and Chapter Ten, "Constructing Choice-Type Items."

Procedures for establishing and measuring reliability are included.


\textsuperscript{37}Ibid., pp. 436-7.

\textsuperscript{38}Ibid., pp. 359-60.

Fremer's (1973) Planning a Test - A Filmstrip Unit on Basic Measurement Principles was designed to cover the basics on measurement concepts.

Test Construction by Wood (1960) was recommended highly by Lyman (1963). Lyman stated:

A thorough but concise and readable job -- as only this author could do it! 39

Lyman's (1963) Test Scores and What They Mean is a popular and highly readable book for the layman. The focus is on standardized tests.

Standards for Educational and Psychological Tests by the American Psychological Association is necessary reading for anyone depending on standardized tests. Many of the standards can be applied to teacher-made tests.

---

3. Guides for Student Evaluation of Instruction: Costin, Greenough, and Manages (1971) performed an ambitious review of literature on undergraduate student evaluation of instruction. They stated:

"The purpose of this article is to review extensively and critically empirical findings concerned with the reliability, validity, and usefulness of student ratings."

Reliability

Studies cited by Costin, Greenough and Manages revealed that instruments in which the classroom instructor was evaluated by students yielded r's from .67 to .89. Studies of internal consistency generally yielded high correlations (.93 for the Illinois Evaluation Questionnaire, for example). "It would appear, then, that students can rate classroom instruction with a reasonable degree of reliability".

Validity

Costin, Greenough and Manages came to the following conclusions concerning validity of student ratings as a result of their review:

1. Criteria used by students tended to be preparedness of the instructor, clarity, and stimulation of curiosity, rather than entertainment value.


41 Ibid., p. 513.
2. Little relationship has been found between rating and grade received; Frey's study (page 90) supports this, also.

3. Majors sometimes rate higher than non-majors.

4. Those taking an elective sometimes rate the elective higher than those taking the course as a required.

5. Upper-class students sometimes rated courses higher than under-classmen.

6. Experienced or higher ranking instructors usually received higher ratings on their instruction than the less experienced.

7. Student ratings of instruction tended to show a "low positive correlation" with ratings of supervisors and colleagues and with measures of student performance after the instruction.

8. The relationship between ratings and teacher personality traits is uncertain or inconclusive.

Costin, Greenough and Manages summarized their review of student evaluation of teaching as follows:

"In conclusion, we wish to emphasize that student ratings of undergraduate teaching fall short of a complete assessment of an instructor's teaching contribution."42

---

Dwyer (1972) concluded that differences in student ratings of instruction do not seem to be due to teacher age or experience, sex, class size, severity of grading procedures, or the rater's maturity. Dwyer's conclusions were based heavily on the Purdue Rating Scale for Instruction.

Werdell (1967) designed a manual to help students initiate their own student evaluation of teaching. It includes many instruments and some words of wisdom. Are course and teacher evaluations of instruction effective in improving instruction? Werdell (1967) quotes Morsh and Wilder: "there is . . . considerable opinion, but little evidence that student ratings will contribute to instructor improvement." (Morsh and Wilder performed a review of literature on identifying effective instructors).

*Centra (1973) performed an experiment to assess whether or not feedback from students was utilized by instructors to improve instruction. Resulting changes in instruction after feedback were modest, or in Centra's words, "hardly striking."

---


Gillmore (1973) suggests that an instrument filled out by students for use by administrators should contain six or less items that are objectively scored and that have face validity. He claims that reliability is not a problem because multiple measures are already built-in — many raters are used. Such an instrument would tap student feelings toward courses, but not necessarily how much they learned.

Gillmore states that such an instrument should not be used by the administration, if at all, by itself. Additional types of inputs are necessary. (The Evaluation of Instruction kit, of course, would provide administrators with examples of alternative types of input for evaluation of instruction. Observation of the instructional processes and follow-up data on students are examples of additional sources of input for evaluation of instruction.)

---

Fritz (1972) found that the two assumptions underlying his study were supported:

1) the values of the student will determine the kind of teacher behavior he is likely to respond to, and
2) the direct assessment of student values is relevant to the study of teacher effectiveness.\(^\text{46}\)

Sanders and Lynch (1973) criticized the practice of generating items on student evaluation instruments that:

1) Do not reflect individual differences in student expectations.
2) Do not have reference points or criteria.
3) Generalizations should not be made across teachers or classes without knowledge of student expectations and criteria.

Aleamoni (1973) stated:

Of the various systems developed for student evaluation of course and instructor, the Illinois Course Evaluation Questionnaire (ICEQ) has perhaps the most extensive reliability and validity data to support it as well as the most extensive norm data base.\(^\text{47}\)


A purpose of the ICEQ is to aid in the improvement of instruction. The instrument was intended to be diagnostic. Frey* (1973) factor analyzed eighteen items on a student evaluation of a teaching instrument and found that six items remained:

a) work-load
b) student accomplishment
c) organization-planning
d) grading
e) teacher's presentations and
f) teacher accessibility

Student ratings on their accomplishment (in calculus, for instance) seemed to correlate highly (+.87) with their achievement on problem-solving examinations.

*Peter W. Frey. Student Instructional Ratings and Faculty Performance. (Paper presented at the annual meeting of the AERA February, 1973.)
Review of Alternatives Summation

The review of literature on Alternatives to Evaluation Kit highlighted alternative frameworks and instruments that could be used in lieu of the Evaluation of Instruction kit; the review also listed guides that might be useful in conjunction with the kit. The most outstanding sources were generally listed first in each section. In order to save the reader the task of hunting back through the review, some of the most useful sources related to evaluation of instruction will be outlined below:

1. Alternative Evaluation Frameworks


2. Complementary Guides

Training packages

The author has not received the training in most of the training packages mentioned and therefore cannot make a comparative statement of worth.

Guides for an overview of evaluation


Guides for performing evaluation


Isaac, Stephen and Michael, William B. 
Handbook in Research and Evaluation. 
San Diego, California: Robert R. 

Sheier, Elaine. Learning 100 Evaluation 
Manual, Eric # ED047192, 1969 
(Microfiche).

3. Alternative Evaluation Instruments

Compendia of listings and reviews

Lake, Dale G.; Miles, Matthew B.; and 
Earle, Ralph B., Jr., eds. Measuring Human Behavior. New York: 

Buros, Oscar Krisen, ed. The Seventh 
Mental Measurements Yearbook. 
Highland Park, New Jersey: The 
Compendia of evaluation instruments


4. Complementary Instrument Construction Guides

Guides to constructing objectives


Guides for evaluation of students


Guides for student evaluation of instruction

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>III. PROCEDURES</td>
<td>97</td>
</tr>
<tr>
<td>A. General Statement</td>
<td></td>
</tr>
<tr>
<td>B. Participants for Study</td>
<td></td>
</tr>
<tr>
<td>C. Interim Validation Overview</td>
<td></td>
</tr>
<tr>
<td>D. Data Collection Framework for Evaluation of Workshop</td>
<td></td>
</tr>
<tr>
<td>E. Chronology of Assessment Activities</td>
<td></td>
</tr>
<tr>
<td>F. Threats to Validity of Evaluation Design</td>
<td></td>
</tr>
<tr>
<td>G. Reliability and Validity of Measures</td>
<td></td>
</tr>
<tr>
<td>H. Data Analysis By the Kit Developer</td>
<td></td>
</tr>
<tr>
<td>IV. RESULTS BY THE KIT DEVELOPER</td>
<td>128</td>
</tr>
<tr>
<td>A. Overall</td>
<td></td>
</tr>
<tr>
<td>B. Interim Validity of Kit</td>
<td></td>
</tr>
<tr>
<td>C. Actual Use of Kit Without Workshop</td>
<td></td>
</tr>
<tr>
<td>D. Workshop Impact</td>
<td></td>
</tr>
<tr>
<td>E. Demonstration of Kit</td>
<td></td>
</tr>
<tr>
<td>F. Interim Validation Demonstration</td>
<td></td>
</tr>
<tr>
<td>V. IMPACT EVALUATION REPORT</td>
<td>147</td>
</tr>
<tr>
<td>VI. INTERPRETATIONS BY THE KIT DEVELOPER</td>
<td>165</td>
</tr>
<tr>
<td>A. Conclusions</td>
<td></td>
</tr>
<tr>
<td>B. Recommendations</td>
<td></td>
</tr>
<tr>
<td>LIST OF WORKS CITED</td>
<td>178</td>
</tr>
</tbody>
</table>
Chapter III

PROCEDURES

General Statement

The procedures build on the kit and workshop that were developed under the first three objectives of the study. The procedures were designed to assess the workshop, and more centrally, the kit. In addition, the procedures were meant to serve as an example, a demonstration. The contents of the kit were used to evaluate the workshop. Using the kit in this way was meant to demonstrate the kit's utility. The interim validation procedures, on the other hand, were used as an example of early assessment of educational products, such as kits.

Overall, the procedures in this study were designed to address the following objectives:

Objective #4: To evaluate the impact of the Evaluation of Instruction Workshop.

Objective #5: To demonstrate the utility of the Evaluation of Instruction kit through its application in evaluating the Evaluation of Instruction Workshop.

Objective #6: To assess the actual use of the Evaluation of Instruction kit without training potential
users.

Objective #7: To assess the interim validity of the Evaluation of Instruction kit.

Objective #8: To demonstrate the application of a potentially transferable method of interim validation of an educational product.

The procedures section is composed of seven concise parts: 1) a description of participants in the study; 2) an overview of how the entire study fits together in terms of interim validation; 3) an outline of the data collection framework used for evaluating the workshop; 4) a chronological description of the assessment activities used in the study; 5) a discussion of threats to the validity of the evaluation design; 6) results on the validity and reliability of instruments used to assess implementation of the workshop; and 7) data analysis by the developer of the kit. Analysis, results, and a judgement by the impact evaluator, however, are a central part of the design and follow the Results By the Kit Developer section, page128.
Participants for Study

The participants included instructors, instructional developers, and administrators in the field of developmental disabilities. They were concerned with higher education or continuing education to the extent that they were willing to commit three days to attend an evaluation of instruction workshop.

A comparison between the nine workshop participants and eight of the mail-out recipients is included in Figure 7.

<table>
<thead>
<tr>
<th>BACKGROUND COMPARISONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CATEGORY</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Employer:</td>
</tr>
<tr>
<td>State Dept.</td>
</tr>
<tr>
<td>College or Univ.</td>
</tr>
<tr>
<td>State School</td>
</tr>
<tr>
<td>Research/Dev. Lab.</td>
</tr>
<tr>
<td>County HR Board</td>
</tr>
<tr>
<td>County HR Program</td>
</tr>
<tr>
<td>Principal Role:</td>
</tr>
<tr>
<td>Administrator</td>
</tr>
<tr>
<td>Evaluator</td>
</tr>
<tr>
<td>Instructor</td>
</tr>
<tr>
<td>Developer</td>
</tr>
<tr>
<td>Speech Therapist</td>
</tr>
<tr>
<td>Consultant</td>
</tr>
<tr>
<td>Coord of Inservice</td>
</tr>
<tr>
<td>Workshop Supervisor</td>
</tr>
<tr>
<td>Highest Degree:</td>
</tr>
<tr>
<td>None</td>
</tr>
<tr>
<td>Diploma, R.N.</td>
</tr>
<tr>
<td>Bachelor's</td>
</tr>
<tr>
<td>Master's</td>
</tr>
<tr>
<td>Doctor's</td>
</tr>
</tbody>
</table>

*those who did not attend at least 2/3 of the sessions
**only 8 of 19 returned background information

FIGURE 7
Participants for the interim validation study were selected in the following way. Potential participants were sent: a) a letter describing the workshop, b) an application form and c) a needs assessment instrument (See Appendix A). Potential participants in Ohio included:

80 Administrators working in the area of developmental disabilities throughout the state

6 Superintendents of state schools

The entire faculties of Educational Development, Exceptional Children, and Curriculum and Foundations at the Ohio State University

The entire faculty of the Nisonger Center, Columbus, Ohio

Employees of any of the above

Respondents to the letter were classified according to whether they were interested in evaluation of instruction or program evaluation or both. Since the majority of responses were addressed to evaluation of instruction, evaluation of instruction was chosen as the topic of the workshop.

The dates that received the most checks (June 11, 12, and 13) were chosen for convening the workshop.

Data were collected from the needs assessment instrument to allow the creation of two groups:

1) WORKSHOP PARTICIPANTS: respondents involved in evaluation of instruction and available to attend the workshop on June 11, 12, and 13.
2) MAIL-OUT RECIPIENTS: respondents involved in evaluation of instruction but unavailable on June 11, 12, and 13; they received the Evaluation of Instruction Kit alone.

As must be expected in the presentation of a workshop for mature professionals, much shifting of plans occurred after the initial letter of notice on the workshop was sent, and many within each group needed to be changed to a different group before the workshop began.

An attempt was made to allow all respondents to attend the workshop who were able to change their commitments and make themselves available on June 11, 12, and 13. Even respondents failing to state that they were involved in evaluation of instruction were telephoned and reinvited. A few (four) respondents were not reinvited because their application indicated that they would be unlikely to benefit appreciably from the workshop because of their primary professional roles.

Interim Validation Overview

The questions posed in this study might best be shown in relationship to each other through a flow chart. Figure 8 shows an overall conceptualization of the study.
FIGURE 8
INTERIM VALIDATION FLOW CHART

Developing Product? → NO → Why Do Study?

Antecedent Necessary? → NO

Antecedent Successfully Implemented?

Recycle → NO

Were Main Sources of Validity Examined?

Recycle → NO

Were Results Positive Enough to Continue Development?

Discontinue → NO

Continue Development
The components of the dissertation can be viewed in a somewhat simplified fashion using the flow chart. Figure 9 show how the more transferable questions can be filled in with the components of this study. Objectives 1-3 formed the foundation for the study. Figure 9 displays the context for objectives 4-7; the flow chart and the entire dissertation were designed to meet objective 8. 

Data Collection Framework for Assessing Kit Utility

Figure 10 displays the source of validity and instrument used to collect data on utility. The most important source of data is actual use of kit. Although much of the information on use of the kit was collected in a telephone interview, workshop participants who claimed to have used the kit were also requested to submit work they had done. Participants submitted evaluation instruments they had constructed and described evaluation designs they were using to substantiate use of the kit.
FIGURE 9
EVALUATION OF INSTRUCTION
INTERIM VALIDATION FLOW CHART

Resource List? Developing Kit? Workshop?

NO

Training Necessary?

YES

Training Successfully Implemented?

YES

Were Main Sources of Validity Examined?

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

NO

Were Results Positive?

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discontinue Mail-out as a dissemination option at present

YES

Continue Development of Kit
FIGURE 10

DATA COLLECTION FRAMEWORK FOR ASSESSING KIT UTILITY
(INTERIM VALIDITY)

<table>
<thead>
<tr>
<th>VALIDITY SOURCE</th>
<th>INSTRUMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Use of Kit*</td>
<td>Telephone Workshop Interview**</td>
</tr>
<tr>
<td>Actual Use of Kit</td>
<td>Telephone Interview for Kit Recipients***</td>
</tr>
<tr>
<td>Perceived Utility of Kit</td>
<td>Product Rating</td>
</tr>
<tr>
<td>Projected Utility of Kit</td>
<td>Rating Evaluation of Instruction</td>
</tr>
<tr>
<td>Projected Utility of Kit Instruments</td>
<td>Evaluation Products</td>
</tr>
<tr>
<td>Demonstrated Utility of Kit</td>
<td>None; Figure 11 Contains Documentation of Use</td>
</tr>
</tbody>
</table>

* Products by participants were used to validate self-report.

** Used with workshop participants 3-5 weeks after workshop.

*** Used with mail-out recipients 3-5 weeks after workshop.
Data Collection Framework for Evaluation of Workshop

The data collection framework for the workshop closely follows the framework presented in the Evaluation of Instruction kit. Using the framework of the kit, the data collection framework is displayed in Figure 11.

The evaluation design for the workshop was meant to meet two objectives: a.) to evaluate the impact of the Evaluation of Instruction Workshop and b.) to demonstrate the utility of the Evaluation of Instruction kit through its application in evaluating the Evaluation of Instruction Workshop.

Instrument Selection

Instruments were selected from the kit for the workshop for eight reasons:

1) Build the workshop partly around the specific group of participants. The instruments in this category include: **Entry Achievement** and **Background Information**.

2) Assess pre-post impact. The instruments in this category include: **Entry Achievement**, **Entry Skill**, **Entry Knowledge**, **Entry Attitude**, **Exit Achievement**, **Exit Skill**, **Exit Knowledge**, **Exit Attitude** and **Achievement Criteria**.

3) Adjust the workshop as it progressed. The instruments in this category include: **Gestalt Statements** and **Criteria of Participants**.

4) Expose participants to a wide range of instruments. All of the instruments fit this category, but the **Group Effectiveness** instrument served this purpose only.
5) Assess the impact of the workshop in terms of judgments of participants toward the workshop. Instruments in this category include: Workshop Outcomes, Gestalt Statements, Debriefing Form, and although they were not "instruments", the Evaluation Reports of participants.

6) Gain diagnostic information to revise the workshop for the next presentation. Instruments in this category include: Gestalt Statements, Debriefing Form, Topic Feedback, Workshop Outcomes, and "Evaluation Reports".

7) Gain information on perceptions of participants toward the kit. The instrument used for this purpose was the Kit Rating instrument.

8) Gain information on the longer term impact of the workshop on participants. The Follow-up Achievement and Follow-up Interview Instruments fit this category.

See Appendices for copies of instruments; locations by page number are located under the listing of each instrument in Figure: 11.

The procedures so far mentioned were intended to acquaint the reader with the components of the study. The next section of the dissertation is intended to provide the reader with a clearer perspective of the sequence of events related to assessment.
**FIGURE 11**

DATA COLLECTION FRAMEWORK FOR WORKSHOP

<table>
<thead>
<tr>
<th>BEFORE WORKSHOP</th>
<th>BEGINNING WORKSHOP</th>
<th>THROUGHOUT WORKSHOP</th>
<th>END OF WORKSHOP</th>
<th>AFTER THE WORKSHOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of Materials:</td>
<td>Entry Skill:</td>
<td>Gestalt Statements:</td>
<td>Evaluation Reports:</td>
<td>Follow-Up</td>
</tr>
<tr>
<td>(By a specialist in instructional technology): workshop agenda was analyzed</td>
<td>Demonstrated competency level to perform evaluation of instruction; pg. 209</td>
<td>Open-ended instrument to collect reactions of participants to workshop experience; pg. 221</td>
<td>Products by two groups of participants developed to assess workshop; pg. 235</td>
<td>Interview:</td>
</tr>
<tr>
<td>Entry Achievement:</td>
<td>Entry Knowledge:</td>
<td>Goals for Session:</td>
<td>Debriefing:</td>
<td>Telephone interview to find how many participants used kit 3-5 weeks after workshop; pg. 252</td>
</tr>
<tr>
<td>Current self-reported competency level to do evaluation; pg. 193</td>
<td>Demonstrated comprehension of evaluation of instruction information; pg. 208</td>
<td>Perceived attainment by participants of workshop objectives; pg. 223</td>
<td>Responses of participants to workshop in an interview format; pg. 226</td>
<td></td>
</tr>
<tr>
<td>Participant Expectations on Achievement:</td>
<td>Participant Criteria:</td>
<td>Topic Feedback:</td>
<td>Workshop Outcomes:</td>
<td>Follow-Up Achievement:</td>
</tr>
<tr>
<td>Expected competency level to do evaluation of instruction at end of workshop; pg. 207</td>
<td>Designated priorities of participants in judging workshop experience; pg. 212</td>
<td>Response of participants to topics of workshop; pg. 222</td>
<td>Descriptive statements of workshop participants on their workshop experience, includes statement and rating of objectives attained; a &quot;goal free&quot; instrument; pg. 232</td>
<td>Same as Entry Achievement; pg. 250</td>
</tr>
<tr>
<td>Participant Criteria:</td>
<td>Entry Attitude:</td>
<td>Instrument Critiquing:</td>
<td>Exit Skill: Same as Entry Skill; pg. 229</td>
<td></td>
</tr>
<tr>
<td>Designated priorities of participants in judging workshop experience; pg. 212</td>
<td>Opinionnaire on evaluation; pg. 210</td>
<td>Participants and workshop presenters critiqued instruments by participants.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background Information: pg. 215</td>
<td></td>
<td></td>
<td>Instrument Rating: 2</td>
<td></td>
</tr>
</tbody>
</table>

specialists compared the quality of instruments constructed by participants after the workshop with instruments constructed before the workshop.
<table>
<thead>
<tr>
<th>BEFORE WORKSHOP</th>
<th>BEGINNING WORKSHOP</th>
<th>THROUGHOUT WORKSHOP</th>
<th>END OF WORKSHOP</th>
<th>AFTER WORKSHOP</th>
</tr>
</thead>
</table>

Exit Knowledge:
- same as Entry Knowledge; pg. 227

Exit Attitude:
- same as Entry Attitude; pg. 230

Exit Achievement:
- same as Entry Achievement; pg. 228

Product Rating:
- Perceived Utility of kit; pg. 233
Chronology of Assessment Activities

Mail-out respondents and workshop participants received a registration application form and a self-rating achievement form to fill out and return. They received the forms approximately three weeks before the beginning of the workshop (see Appendix A1). During the same period an instructional technologist assessed the workshop materials.

Workshop participants were assessed shortly after the beginning of the Evaluation of Instruction Workshop. They were given a packet of instruments to fill out (see Appendix B1). The instruments included:

- Background Information
- Entry Skill
- Entry Knowledge
- Entry Attitude
- Criteria of Participants
- Achievement Criteria

In the early afternoon of the first day, the participants were asked to fill out the *Gestalt Statements* instrument. They were told that they were free to answer it in any way they wished.

At the end of the first day of the workshop the participants were asked to fill out the *Topic Feedback* instrument; they were told to consider the entire day
as a session when they saw the word "session" on the instrument.

Mid-morning of the second day the participants were again asked to fill out the Gestalt Statements instrument as before.

The second day ended with participants filling out a Goals for Session instrument.

During the afternoon of the third day the participants divided into two groups (teachers in one, program evaluators in the other). Each group was given a typed copy of the raw data from instruments used to assess and monitor the workshop. They used the data to construct an evaluation report on the workshop. The developer of the workshop then used the Debriefing Form to interview the participants.

The final activity of the third day involved assessment of the workshop and the kit. Participants responded to the following instruments:

Exit Achievement
Exit Skill
Exit Knowledge
Exit Attitude
Rating Evaluation of Instruction Kit
Workshop Outcomes
The mail-out respondents were sent the following instruments to complete and return on the day the workshop began:

- Background Information
- Entry Skill
- Entry Knowledge
- Entry Attitude

A letter was enclosed with the instruments (see Appendix A2).

One and one-half weeks after the end of the workshop, both workshop participants and mail-out recipients were sent a letter. Workshop participants received a Follow-Up Achievement instrument with their letter (see Appendix E1). Mail-out recipients received a Follow-Up Achievement instrument, a Rating Evaluation of Instruction Kit instrument, and a Follow-Up Attitude instrument with their letter (see Appendix E2).
All Workshop Participants and Mail-out Recipients were given a telephone interview by an impact evaluator three to five weeks after the end of the workshop. They were informed that all evaluation designs and instruments that they had developed since the end of the workshop were needed very much. They were told that their work would be critiqued and returned if they would send it to Jerry Adams. (See Appendix E.)

The content validity of the kit was assessed by ratings on the kit (in manual form) by

a) an expert in educational curriculum,
b) an expert in educational innovation and change,
c) an expert in educational dissemination, and
d) an expert in evaluation.

In addition, two evaluation specialists rated the instruments within the kit and the instruments constructed by workshop participants.
Threats to Validity of Evaluation Design

Since the study deals with a comparison of two non-equivalent groups, the internal validity of the study can be challenged; however, the two groups were similar* upon entry on the variable thought to be most important, entry achievement (self-reported) relative to workshop objectives.

Averages of self-report data on a 5-point** anchored scale suggested that the two groups were also similar in terms of experience and training in evaluation. The workshop group's average score on experience was 2.4, while the mail-out respondents' average score was 2.5; both groups had an average score of 2.5 on training in evaluation. The main visible difference between the two groups consisted of their availability or unavailability to attend the workshop.

Since all members of both groups were willing to attend a three-day workshop and also reported need for evaluation activities during June, motivation for using the kit is seen as roughly equivalent between the two groups.

*Rank Sum Test on entry achievement (self-reported)
**1=no experience or training in evaluation
Perhaps the greatest threat to validity in the evaluation procedures was for mail-out recipients. They were offered the workshop and then informed that the workshop would not be held on the date they could attend. Some recipients of the kit may have felt bitter about the situation and not used the kit because they could not attend the workshop. Little control against this problem is readily apparent: if members of the group are not offered the workshop, then the investigator would not know if they were motivated enough to take off three days. The present approach was a trade-off designed to maximize the likelihood of somewhat equal pre-treatment motivation on the part of those selected as workshop participants and those selected for the "mail-out" group. In order to minimize the negative reactions or reluctance in the mail-out recipients the investigator offered consultation services in lieu of the workshop and contingent upon use of the kit.

Two potential threats to validity for workshop participants include "halo" effects, generalized responses after superficial impressions, and Hawthorne effects, differential behavior due to knowledge of participating in a study. Effective control
groups are generally intended to handle these effects in an experimental design. However, successive replications seem a viable alternative for this evaluation design. The workshop might be repeated outside the context of a study to control for Hawthorne effects. "Halo" effects, on the other hand, seem to be an inherent part of workshops; repeated presentations might be the best control for intensity of the effects.
Reliability and Validity of Measures

Ultimately several validation studies will be necessary for precise discriminations to be made concerning the impact of the Evaluation of Instruction Kit.* This study was designed primarily to gauge receptivity to the kit with two modes of dissemination. Some data were collected on the measurement devices, however.

Instrument for Rating Kit (by Experts)

A rating scale was developed to collect judgements of experts concerning the content validity of the kit. Items for the instrument were derived from categories isolated by Ronald G. Havelock (1969) in his review of literature concerning essential ingredients for a successful educational innovation. The instrument was critiqued by an evaluation specialist** and then revised; it was then critiqued by a potential user*** of the kit and revised again.

Instrument for Rating Kit (by Users and Potential Users)

A rating scale was developed to collect judgements by users and potential users on the perceived utility of

*Stepwise evaluation procedures have also been advocated by Smock and Crooks (1973) and by Rossi (1972). They asserted that global or "rough" assessment should precede more discriminating procedures.

**Ms. Kay Adams

***Ms. Pat Kennedy
of the kit. The instrument was critiqued by an evaluation specialist* and then revised.

**Interview Schedule**

The interview schedule was generated by the developer of the workshop and then pilot tested and revised by the impact evaluator. The schedule was developed to collect data regarding the central questions of the study:

Was the kit used?

How much?

In what ways?

The schedule was piloted on two participants who did not remain in the workshop long enough to be counted as "official" participants and on two who received the kit in the mail too late to be counted in the "mail-out" group. The interview schedule was revised and interviews were conducted three weeks after the workshop was completed. Interview items were based on objectives of the study.

**Workshop Impact Assessment Instruments**

1. Evaluation Opinionnaire

The opinionnaire was constructed by a nationally recognized evaluation leader, Blaine Worthen, and his

*Dr. William E. Loadman
assistant, Marilyn Averill. The instrument was used to assess impact of an evaluation workshop held by Daniel Stufflebeam and Michael Scriven. Pre-post analysis by Worthen and Averill revealed that the opinionnaire was non-reactive.

In addition, the Item Analysis program (for Likert scales) by Joyce Johnson and June McCabe was run on the items of the opinionnaire to gauge its reliability. A Kuder-Richardson Scale Reliability Coefficient of .64 was obtained.

2. Criterion Referenced Instruments

The instruments used to assess knowledge, skill, and achievement were criterion-referenced. The instruments were derived directly from the objectives of the workshop and were judged by two evaluation specialists to have "mediocre" to "good" content validity (see page 159).

*Joyce Johnson and June McCabe, Item Analysis C6.01.049, Data Center, College of Administrative Science, The Ohio State University, August 1, 1972.
Congruence reliability is high when multiple measures are used which have high inter-correlations. Workshop participants should score similarly on all three instruments if the instruments reliably measure degree of objective attainment.

Since the achievement instrument was the easiest of the three to administer and analyze, Kendall's $K$ statistic and estimate of association were used to determine the degree of association between skill and achievement results and between knowledge and achievement results.

Entry knowledge and achievement were positively associated ($\alpha = .017$) and the estimate of the degree of association was high ($\hat{\eta} = .68$). Entry skill scores on the nine workshop participants were too low* to allow effective correlation since six out of nine had "zero" totals.

Exit skill and achievement were positively associated ($\alpha = .022$) and the estimate of the degree of association was high ($\hat{\eta} = .55$). Exit knowledge scores were too high* to allow effective correlation since five out of nine had the highest possible score.

The analysis results suggest that participants have some ability in estimating their achievement level. Further analysis using more participants and inferential procedures should precede using self-report on achievement as the sole indicator of success. The skill and knowledge instruments might be helpful as adjuncts in the instructional process and retained for that reason as well.

*Kendall's $K$ statistic and estimate is computed using ranking. When a large number of the scores are tied, the procedure is
3. Workshop Outcomes: Instrument

A "goal free" instrument was developed and pilot tested on a set of participants similar to those in the study. The subjective results were judged by the workshop developer to yield useful indications on outcomes; the instrument and results had high face validity for the developer.

4. Formative Assessment Instruments

The Topic Feedback instrument was used by the developer at the end of over thirty workshop presentations. The instrument was seen as useful to the developer of the workshop for collecting information to revise the workshop. The developer used the Gestalt Statements instruments in six workshops; the instrument was judged as useful for collecting information to monitor the workshop activities. The Goals for Session instrument was a modification of an instrument developed to assess another workshop on evaluation. The developer judged that the instrument would be useful in collecting information on perceptions of participants regarding attainment of some workshop goals. The results were to be used as a basis for revising future sessions. The Criteria for Participants instrument was a modification of an instrument developed by Ms. Helen Armstrong to collect information on a set of continuing education workshops. The
items were judged as useful by the developer of the workshop in collecting information to monitor discrepancies between relevant values of participants and the presentor.

The Analysis of Materials instrument was pilot tested on a workshop developer,* two secretaries,** an evaluation specialist,*** and an instructional technologist.**** Most of the items were judged as useful by the developer of the workshop in collecting information to revise the workshop before the workshop began. The instrument was adapted from an instrument developed by Eash(1970).*****

All of the above formative instruments were selected or developed on the basis of face validity for the developer of the workshop.

Workshop Instruments: Content Validity Overall

Two evaluation specialists rated twelve of the instruments used to assess the workshop. One specialist

*Mr. John Bendekovic

**Ms. Carolyn Moss and Ms. Nancy Rutter

***Ms. Kay Adams

****Ms. Barbara Janko

rated eight of the instruments as good, three as mediocre, and one as poor in terms of content validity (see Appendix F, pp. 266). The second specialist rated eight instruments as good, one as mediocre and left three unrated since they addressed objectives for the workshop developer rather than objectives for the workshop. The specialists agreed on four instruments, and had unknown agreement on the three instruments unrated by one specialist.
Summary

It may be noted that face validity and content validity were heavily emphasized in this study. Smock and Crooks (1973)* asserted that when evaluation of instruction is focused on formative evaluation for the instructor, face validity may frequently be the most important type of validity for an instrument. Bloom noted the emphasis on content validity in evaluation as contrasted to other types of validity:

As we review these different types of validity, it is striking that each of the testing approaches has emphasized particular types of validity and has developed techniques and procedures for utilizing the preferred types of validation.

Thus, evaluation has stressed content validity . . .**

Bloom also advocated emphasizing content and construct validity before attempting to establish concurrent or predictive validity.***


***Ibid., p. 40.
The results on sixteen of the measures in this study have high congruence reliability as explained by Bloom.* That is, most of them suggest the same conclusion: dissemination of the kit via the workshop was successful and dissemination via the mail was not.

---

*Ibid., 42-3.
Data Analysis By the Kit Developer

The purpose of the study was to assess the validity of the Evaluation of Instruction kit via two modes of dissemination. Since the per cent of individuals who used the kit in each mode was so radically different (0% and 75%), statistical analysis on use seemed unwarranted. Use of the kit was the most important indicator of interim validity.

A Wilcoxon Signed Rank test was used to make the following less important comparisons:

Workshop Participants

Achievement: pre-workshop vs. post-workshop
post-workshop vs. follow-up
expectation vs. post-workshop
expectation vs. follow-up

Skill: pre-workshop vs. post-workshop

Knowledge: pre-workshop vs. post-workshop

Opinions: pre-workshop vs. post-workshop

Equivalence between workshop participants and mail-out recipients was also tested. A Rank Sum test was used to compare the self-rated entry achievement level of the eight workshop participants and the six mail-out recipients who participated in the follow-up interviews.
Three of the six kit recipients interviewed did not submit entry information other than entry achievement so no other entry level comparisons between groups was feasible.
Chapter IV
RESULTS BY THE KIT DEVELOPER

Overall
Results on the impact of the workshop that are present by the developer of the workshop and kit must be viewed as subject to bias; the presentation is made for a cohesive demonstration of how the Evaluation of Instruction kit was applied to the workshop and to cohesively demonstrate the overall application of the interim validation procedures. A less biased and less complete treatment of the workshop results are found in the Impact Evaluation Report (pp.147-164). More complete and less biased results on the interim validity (usefulness) of the kit are found in the Impact Evaluation Report also.

The results of the first three objectives were educational products--a kit and a workshop. The information in the "Review of Literature" formed a basis for the development of the Evaluation of Instruction kit. The Evaluation of Instruction workshop was developed to help assess the worth of the kit. All of the first three objectives were therefore met--a product was developed
for each objective. Objectives #4, #6, and #7 were designed to assess the quality of the first three products, especially the Evaluation of Instruction kit. Objectives #5 and #8 were designed to provide a demonstration of some evaluation procedures.
Results By Objectives

Figure 12 displays the results of objectives #4-#8 in summary form. The objectives are listed in order of importance to the writer. The meaning of the different categories in the Figure is as follows:

The dependent variables are the areas of change expected with each objective. The indicators are more specific statements of intended outcomes for each objective. The next category indicates when information was collected to find if the desired outcome was achieved. The instruments cited in the Figure are only generic, not specific. The "N" stands for the number of individuals who were sources of information. The sources of information are the groups of individuals who were assessed or asked regarding intended outcomes for each objective. The data collectors are individuals administering instruments or conducting interviews. The results summary is an overview of the most important outcomes of the study from the point of view of the kit developer.

A narrative summary of results follows Figure 12. The narrative summary is written in terms of objectives and dependent variables for each objective. The reader is reminded that the narrative summary is subject to bias and is presented for purposes of demonstration.
<table>
<thead>
<tr>
<th>Objectives</th>
<th>Dependent Variables</th>
<th>Indicators</th>
<th>Time Data Collected</th>
<th>Instrument Used</th>
<th>N</th>
<th>Source of Information</th>
<th>Data Collector</th>
<th>Summary Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#7: Assess interim validity of kit</td>
<td>actual use of kit</td>
<td># of wsp* using kit</td>
<td>3-5 wks after workshop</td>
<td>interview schedule</td>
<td>8</td>
<td>workshop participants</td>
<td>Armstrong</td>
<td>75% used kit</td>
</tr>
<tr>
<td></td>
<td>perceived utility of kit</td>
<td>ratings of kit</td>
<td>end of workshop</td>
<td>rating form</td>
<td>9</td>
<td>workshop participants</td>
<td>Adams</td>
<td>4.05 average on 1-5; 1=low</td>
</tr>
<tr>
<td></td>
<td>projected utility of kit</td>
<td>ratings of kit</td>
<td>after workshop</td>
<td>rating form</td>
<td>5</td>
<td>experts</td>
<td>Adams</td>
<td>3.16 average on 1-5; 1=low</td>
</tr>
<tr>
<td></td>
<td>projected utility of instruments in kit</td>
<td>ratings of instruments in kit</td>
<td>after workshop</td>
<td>rating form</td>
<td>2</td>
<td>evaluation specialists</td>
<td>Adams</td>
<td>mostly medium to good</td>
</tr>
<tr>
<td>#6: Assess use of kit w/o training</td>
<td>use of kit</td>
<td># of mor** using kit</td>
<td>3-5 wks after workshop</td>
<td>interview schedule</td>
<td>6</td>
<td>mail-out recipients</td>
<td>Armstrong</td>
<td>0% used kit</td>
</tr>
<tr>
<td>#4: Evaluate workshop impact</td>
<td>skill</td>
<td>sig. pre-post gain</td>
<td>start &amp; end of workshop</td>
<td>criterion referenced</td>
<td>9</td>
<td>workshop participants</td>
<td>Adams</td>
<td>sig. at &lt; .004</td>
</tr>
<tr>
<td></td>
<td>knowledge</td>
<td>sig. pre-post gain</td>
<td>start &amp; end of workshop</td>
<td>criterion referenced</td>
<td>9</td>
<td>workshop participants</td>
<td>Adams</td>
<td>sig. at &lt; .004</td>
</tr>
<tr>
<td></td>
<td>achievement</td>
<td>sig. pre-post gain</td>
<td>start &amp; end of workshop</td>
<td>self-rating form</td>
<td>9</td>
<td>workshop participants</td>
<td>Adams</td>
<td>sig. at &lt; .02</td>
</tr>
<tr>
<td></td>
<td>attitude</td>
<td>sig. pre-post gain</td>
<td>start &amp; end of workshop</td>
<td>opinionnaire</td>
<td>9</td>
<td>workshop participants</td>
<td>Adams</td>
<td>not sig.</td>
</tr>
<tr>
<td>actual use (see #7)</td>
<td>seeking help</td>
<td># of wsp* seeking help</td>
<td>2-3 wks after workshop</td>
<td>interview schedule</td>
<td>8</td>
<td>workshop participants</td>
<td>Armstrong</td>
<td>4 sought help</td>
</tr>
</tbody>
</table>

* wsp = workshop participants
** mor = mail-out recipients

FIGURE 12
<table>
<thead>
<tr>
<th>Objectives</th>
<th>Dependent Variables</th>
<th>Indicators</th>
<th>Time Data Collected</th>
<th>Instrument Used</th>
<th>N</th>
<th>Source of Information</th>
<th>Data Collector</th>
<th>Summary Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4: Continued</td>
<td>instrument quality</td>
<td>categorize by quality</td>
<td>before and after workshop</td>
<td>none</td>
<td>2</td>
<td>evaluation specialists</td>
<td>Adams</td>
<td>no pre-post gain</td>
</tr>
<tr>
<td># of instruments</td>
<td>pre-post</td>
<td>pre-post w/o an instrument</td>
<td>before and after workshop</td>
<td>questionnaire and interview schedule</td>
<td>8</td>
<td>workshop participants</td>
<td>Adams &amp; Armstrong</td>
<td>3 w/o pre; 0 w/o post</td>
</tr>
<tr>
<td># of designs</td>
<td>pre-post</td>
<td>pre-post w/o a design</td>
<td>before and after workshop</td>
<td>questionnaire and interview schedule</td>
<td>8</td>
<td>workshop participants</td>
<td>Adams &amp; Armstrong</td>
<td>7 w/o pre; 2 w/o post</td>
</tr>
<tr>
<td>judgements toward workshop</td>
<td># of (+) vs (-)</td>
<td>first day of workshop</td>
<td>open ended</td>
<td>9 workshop participants</td>
<td>Adams</td>
<td>34 (+); 43 (-)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>comments</td>
<td>second day of workshop</td>
<td>open ended</td>
<td>9 workshop participants</td>
<td>Adams</td>
<td>29 (+); 9 (-)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>summary statement</td>
<td>third day of workshop</td>
<td>&quot;goal free&quot;</td>
<td>9 workshop participants</td>
<td>Adams</td>
<td>main objectives met</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5:* Demonstrate kit utility</td>
<td>application of kit</td>
<td>workshop evaluation framework &amp; instruments</td>
<td>throughout study</td>
<td>20 from kit</td>
<td>9</td>
<td>workshop participants &amp; experts</td>
<td>Adams &amp; Armstrong</td>
<td>all of framework &amp; 20 parts used</td>
</tr>
<tr>
<td>#8: Demonstrate interim validation procedures</td>
<td>application of procedures used for objective #7</td>
<td>procedures</td>
<td>end of &amp; after workshop</td>
<td>interview schedule, 3 rating forms, &amp; objective #5</td>
<td>17</td>
<td>workshop participants, kit developers, experts, specialists</td>
<td>Adams</td>
<td>the procedures were used</td>
</tr>
</tbody>
</table>

*Objective #5 is also part of objective #7.
Interim Validity of Kit

1. Actual Use of Kit

The central purposes of the study were achieved. The kit was found to be useful for at least one of the groups studied. A comparison between groups revealed that 75% of workshop participants used the kit although none of the mail-out recipients used the kit; see Impact Evaluation Report, page 147.

2. Projected Utility of Kit as a Whole

The projected utility of the kit was not so clear. Experts rated the kit "mediocre" to "good" in general, but rated it "poor" in ease to use for the intended audience (see Impact Evaluation Report). Instructions on use of the kit were seen as sparse and choppy.

3. Projected Utility of Kit Instruments

Two evaluation specialists rated the projected utility of instruments in the Evaluation of Instruction kit. Twenty-two instruments were rated. One specialist rated seventeen instruments as good, five as mediocre, and none as poor (see Appendix F, pp. 266-7). The second specialist rated nine as good, eleven as mediocre, and three as poor. The two specialists agreed on ten ratings, disagreed by one on ten ratings, and completely disagreed on two ratings. The two specialists agreed that seven out of the twenty-one instruments were good, agreed on three instruments as mediocre, and they did not agree on any
instruments as poor. In general, then, almost all of the instruments had mediocre to good ratings.

4. Perceived Utility of Kit

The kit was perceived as useful by the workshop participants at the end of the workshop (see Impact Evaluation Report).

Actual Use of Kit Without Workshop

None of the six mail-out recipients contacted in the follow-up telephone interviews reported using or attempting to use the Evaluation of Instruction Kit (see Impact Evaluation Report).

Workshop Impact

1. Pre-Post Impact

Four measures of workshop participants were collected at the beginning of the workshop and again at its close. These were evaluation skill, knowledge, achievement, and opinion toward evaluation. A nonparametric statistic, the Wilcoxon Signed Rank Test was computed, using the pre-post data of each measure to discover if and to what degree the workshop experience affected the participants. (The Wilcoxon has high relative efficiency in analyzing pre-post effects of a group without a normal distribution of scores). The null hypothesis of no workshop effect versus the alternative hypothesis of a positive workshop effect was tested. A Hodges-Lehmann estimate of the amount of difference between
pre-test and post-test scores was then calculated. The results of these tests are presented in Table 1.*

### TABLE 1

Pre-Post Differences of Workshop Participants on Measures of Evaluation Skill, Knowledge, Achievement, and Opinion

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pre-Post Gain</th>
<th>( \alpha )</th>
<th>( T^{**} ) and ( t_{(\alpha, n)} ) value</th>
<th>Estimate of Pre-post total score difference((^{6}))</th>
<th>Maximum Possible Score</th>
<th>( N^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill</td>
<td>Significant</td>
<td>.004</td>
<td>( T^{**}=36 &gt; t(\alpha=0.004, 8) ) =36</td>
<td>14.5 points</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Significant</td>
<td>.004</td>
<td>( T^{**}=36 &gt; t(\alpha=0.004, 8) ) =36</td>
<td>6.25 points</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Achievement</td>
<td>Significant</td>
<td>.02</td>
<td>( T^{**}=40 &gt; t(\alpha=0.02, 9) ) =40</td>
<td>10.5 points</td>
<td>60</td>
<td>9</td>
</tr>
<tr>
<td>Opinion</td>
<td>Not Significant</td>
<td>.098</td>
<td>( T^{**}=28 &lt; t(\alpha=0.055, 8) ) =30</td>
<td>2.5 points</td>
<td>50</td>
<td>8</td>
</tr>
</tbody>
</table>

* The original population consisted of nine (9) participants. When computing the signed rank statistic, \( Z \) values (pre-test score subtracted from post-test score) equal to 0 are discarded and the sample size reduced accordingly.

** \( T^{**} \) is the Wilcoxon Signed Rank value for differences between medians.
The Achievement instrument was administered to workshop participants at four different times to measure: Pre-workshop achievement, Expectation of workshop achievement, Post-workshop achievement, and Follow-up achievement (three to five weeks after the workshop). The Wilcoxon Signed Rank statistic and Hodges-Lehmann Estimate were again computed to test for significant differences between the measures and to estimate the amount of the differences. The results of these tests are presented in Table 2.
### TABLE 2

**Differences Among Four Administrations of the Achievement Instrument**

<table>
<thead>
<tr>
<th>Administration</th>
<th>Gains</th>
<th>$\alpha$</th>
<th>$t^+ and t_{(\alpha,n)}$ values</th>
<th>Estimate of total scores differences* $\hat{e}$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-post</td>
<td>Significant</td>
<td>.02</td>
<td>$t^+ = 40 \geq t(.02, 9)$</td>
<td>14.5</td>
<td>9</td>
</tr>
<tr>
<td>Post-Follow-up</td>
<td>Significant</td>
<td>.008</td>
<td>$t^+ = 28 \geq t(.008, 7) = 28$</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Expectation Follow-up</td>
<td>Not Significant</td>
<td>.234</td>
<td>$t^+ = 19 &lt; t(.055, 7) = 24$</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Expectation Post</td>
<td>Not Significant</td>
<td>&gt; .50</td>
<td>$t^+ = 21 &lt; t(.049, 9) = 37$</td>
<td>1.5</td>
<td>9</td>
</tr>
</tbody>
</table>

*The total possible score for the Achievement Instrument is 60 using the 12-item form of the instrument.*
In summary, pre-post indicators suggested that the workshop had a significant positive effect on evaluation skill, knowledge, and perceptions of achievement of participants. Opinions toward evaluation began as positive and did not change significantly. No significant difference was found between expectation for achievement and post workshop ratings on achievement -- this suggests that the workshop met expectations of participants in that area. Although follow-up ratings on achievement were higher than post-workshop ratings, follow-up ratings did not significantly exceed expectations.
2. Formative Indicators of Impact

Attitude of the participants toward the workshop is another reflection of the implementation of the workshop in addition to learning by participants.* The open-ended Gestalt Statements instrument was distributed twice during the workshop. A total of 34 comments were positive, 43 were negative, and 16 could not be readily placed in either category from the first administration of the instrument. A total of 29 comments were positive, 9 were negative, and 13 could not be readily placed in either category from the second administration of the instrument. The workshop started out slightly negative, it seems, and improved considerably.

Written evaluation reports by participants at the end of the workshop suggested that overall, the workshop went well. The summary statement in one report was:

"In general, a jolly good time was had by all and all felt positive about the experience."

(The second report had no summary statement on overall worth.)

The workshop outcomes instrument administered at the end of the workshop required participants to state the objectives of the workshop that they had experienced. The three most frequently mentioned were: 1) develop instruments, 2) use the model, and 3) learn terminology, respectively. The first objective was seen as met "partially" to "greatly", the second objective was met

* For a diagnostic treatment of the same data, see 143.
"greatly", and the third objective was met "partially". Since most of the workshop objectives revolved around competence in development of instruments, the workshop might be seen as "partially" to "greatly" a success on that criterion if participants are good judges. (Pre-post measures on their skills suggest that they did gain skills but the gain in skills was not readily apparent in the particular participant's instruments rated by experts).

In addition to the data collected during the workshop, the author looked at the judgmental comments of participants collected three to five weeks after the workshop. The following is a list of the comments on the follow-up interview schedule:

1) The workshop was altered to meet my needs;
2) More concerned with evaluation;
3) Realized hidden skill in evaluation;
4) Have incentive to use evaluation;
5) More self-reliant and less reliant on evaluation specialist;
6) Major asset of the workshop was its orientation toward producing a usable product;
7)Entered workshop believing its aim was process evaluation. Developed interest in evaluation (doing reading in area). Reshaped thinking. Used jargon with new employer;
8) Created feeling that need to learn more about evaluation;
9) Workshop helped;
10) Workshop great. Appreciate kit - it provides better reference materials than most lecture notes;

11) Used workshop information to apply for a grant (better design and proposed use of evaluation).

12) Very good. Basis for much future work.

The attitude of participants toward the workshop in the period three to five weeks after the workshop seems to have been very good.

3. Summary

Overall, the workshop seems to have been successfully implemented.

Application of Kit

The Appendices documents show how the framework and instruments in the Evaluation of Instruction kit influenced the evaluation of the Evaluation of Instruction Workshop. Further documentation is offered by the Workshop Impact section (pp. 134-142) and Figure 11 of the dissertation. However, the Appendices, Figure 11, and the Workshop Impact sections fail to reveal how the kit was used to perform diagnostic formative evaluation - evaluation used to make revisions.
1. Formative Diagnostic Results on Workshop

The information collected during the workshop was designed primarily for diagnosis of problems and strengths to help revise the workshop. The Gestalt Statements were helpful in revising the workshop as it progressed. The first set of returns revealed 12 comments of concern about an inability to follow the jargon used in the workshop by the presenter. The workshop was immediately stopped and participants generated a list of terms that they did not understand; the presenter wrote them on the board and defined each; the presenter had the list of definitions typed, duplicated, and presented to participants within an hour of actual workshop time (lunch intervened).

Five comments revealed a desire for more formal instruction. One extra lecture was developed, but, in general, the presenter chose not to change the basic format of the workshop.

When the presenter attempted to divide the participants into homogenous groups, he discovered that backgrounds of participants had been poorly communicated. Several participants were looking for program level evaluation or evaluation of children. The workshop was redesigned to accommodate the problem. Two groups were formed: one group worked on program evaluation as it applied to the kit and the second group worked on
evaluation at the classroom level.

The second administration of the Gestalt Statements instrument revealed that participants were appreciating the "application" orientation of the workshop. Plans to use what was learned upon returning from the workshop were also expressed frequently. Several comments reflected that the presenter had given poor directions concerning activities the participants should engage in during the workshop. The instructor asked for suggestions on what to do about his poor directions and said that he would try harder. Several also felt that they lacked a good foundation in evaluation before entering the workshop and that a good foundation would have been very helpful.

The Topic Feedback instrument also revealed that participants were concerned with not understanding the jargon. In addition, however, the responses suggested that the topic of "needs assessment" may have needed expansion. (Since the follow-up study concurred by revealing that the majority of instruments constructed by participants were for needs assessment, the next workshop should reflect that area of interest more strongly.)

At the beginning of the third day the presenter gave the participants individually a choice among alternatives for the focus of the rest of the workshop. They almost
unanimously requested a lecture and focus on overall evaluation designs and integrative activities to pull together what they had learned. A lecture on evaluation designs and models was presented; later participants worked on generating an evaluation report on the workshop by using some of the data collected on the workshop.

A debriefing of participants at the end of the workshop revealed that the presenter should take the following action in the next workshop presentation:

a) send basic reference material to participants before the workshop, especially stressing terminology and how to use the kit.

b) start the workshop with heavy emphasis on how to use the kit and how the kit relates to the overall workshop.

c) present an overview of evaluation models early in the workshop.

d) continue the "responsive" approach taken.

Participants evidently thought that the main workshop objectives they experienced were a) to develop instruments, b) to learn terminology, and c) to use an evaluation model. These results on the Workshop Outcomes instrument reveal an amazing oversight by the developer: He never listed "skill in using the Evaluation of Instruction kit" as an objective of the workshop! The third objective, "using
the model," obviously refers to using the kit, however. The specific objective should be added to the next set of workshop objectives.

Since the comments from the follow-up study suggest that workshop participants were pleased with what they had learned, the next workshop seems likely to produce excellent results if the participants' suggestions are heeded and the revisions are made.

Interim Validation Demonstration

An outcome of the entire study is a flow chart demonstrating the writer's conceptualization of interim validation procedures related to utility of a framework or kit. The potentially transferable procedures were portrayed in Figure 8.
Chapter V

IMPACT EVALUATION REPORT

BY

HELEN ARMSTRONG
CONSULTANT EVALUATOR
Preface By Jerry Adams

An impact evaluator was hired to do two things: 1) conduct a telephone follow-up interview with both groups in the study and 2) use all available data for a judgement of worth on the kit.

The report was edited by this writer under the supervision of one of the members of the doctoral committee. The rather large section edited out related to data for revision of the workshop, not impact. The only purpose of the editing was to prevent confusion for the reader and to make the central findings more easily communicated.
TABLE OF CONTENTS

I Purpose
II Evaluation Design
III Limitations
IV Results
   A. Background Information
   B. Kit Rating by Participants in Workshop
   C. Kit Rating by Experts
   D. Kit Use (Telephone Follow-up Survey)
      1. Mail-out Recipients
      2. Workshop Participants
   E. Rating of Evaluation Products
V Conclusions
VI Recommendations
VII Vita of Impact Evaluator
PURPOSE

This evaluation was designed to explore the interim validity of the Evaluation of Instruction Kit and Workshop by studying their impact on utilization of the kit. Differences in effect between those who attended the workshop and those who received only the kit were also explored since the workshop was designed to facilitate use of the kit.

The objectives of the study were prepared by the kit and workshop developer (hereafter called the developer). The objectives of the developer guided the impact evaluation.
EVALUATION DESIGN

Multiple measures were taken for most information needs. This report looks at general congruence reliability resultant from multiple measures. Great reliance was also placed on using instruments and data furnished by the developer. (See data collection framework, page 108).

In addition, a telephone follow-up survey was conducted by the impact evaluator for all workshop participants who had attended at least two-thirds of the sessions and for six kit recipients who had received the kit more than two weeks before the survey. The instrument developed for the survey was field tested on participants who did not meet the preceding criteria and was revised before beginning the actual survey.
LIMITATIONS

The major limitation of this study may have been time. Too little time may have elapsed between the announcement and presentation of the workshop to provide maximum participation. Too little time may also have elapsed between the workshop (and kit receipt) and post evaluation to allow good development, receipt and analysis of resultant instruments. Summer schedules of persons in an educational setting are frequently very different from their schedules during the balance of the year. Vacations -- of either those participating or their co-workers -- may increase demands on participants during the summer. In some instances additional programs (i.e., summer camps) had placed further demands on participants in the survey. As a result, the number of available respondents provided a great deal of data from a very limited audience; however, this data was definitive enough to allow conclusions concerning the differential impact of the two modes of disseminating the kit.

Since the total population (kit recipients and workshop participants) was included in this evaluation, only descriptive reporting is made. There is no opportunity to statistically project the outcome for a larger population; however, some of the results of this evaluation may serve as examples (case studies) for other research.
RESULTS

A. Background Information

Only 17% of all applicants indicated they would be teaching either higher education or continuing education courses in June. This should have signaled to the developer that some alterations needed to be made in either participants or content; however, no evidence exists to indicate response to this need until the first day of the workshop when participants were grouped according to background and an effort was made to individualize the presentation.*

One-third of the workshop participants did not meet the criterion of participating in two of the three day's sessions. Limited data exists on reasons for this; however, data obtained indicated the following:

a. other obligations limited attendance;

b. the workshop was not designed to provide program evaluation information;

c. participants did not expect to gain new information.

This third reason may warrant further investigation. The average percent of items correct (on the pre-test of knowledge) was 62.5% for those who did not stay, compared to 53.3% for those who did stay.

* Note from Developer: Most of the workshop participants responded by mail or by telephone that they would be involved in teaching, developing, or evaluating instruction in June.
B. Kit Rating by Participants in Workshop

Participant's ratings -- according to four criteria
-- follow (a 5-point scale* was used):

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Average Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Ease of understanding</td>
<td>3.6</td>
</tr>
<tr>
<td>(2) Ease of usage</td>
<td>4.1</td>
</tr>
<tr>
<td>(3) Improvement of instruction</td>
<td>4.1</td>
</tr>
<tr>
<td>(4) Planned utilization</td>
<td>4.4</td>
</tr>
</tbody>
</table>

The planned future usage indicated in (4) does, in fact, seem to occur (see Section D).

C. Kit Rating by Experts

The kit was rated by both the workshop participants and by experts in the fields of evaluation, dissemination, innovation, and curriculum, one expert per field. They were asked to rate the kit on a 5-point scale according to the following criteria:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Average Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Importance of concept</td>
<td>3.6</td>
</tr>
<tr>
<td>(2) Importance of information</td>
<td>3.6</td>
</tr>
<tr>
<td>(3) Ease of usage</td>
<td>2.0</td>
</tr>
<tr>
<td>(4) Helpfulness</td>
<td>3.0</td>
</tr>
<tr>
<td>(5) Overall quality</td>
<td>3.4</td>
</tr>
</tbody>
</table>

The low rating for ease of usage differs substantially from the average participant response, yet may well reflect results obtained on mail-out recipients.

* 1 = Low
D. Kit Use (Telephone Follow-Up Survey)

Beginning three weeks after the workshop, telephone survey instruments were field tested on participants who did not meet either the criteria of attending 2/3 of the workshop or criterion of having received the kit at least 2 weeks prior to that time. The instruments were then revised and the actual survey was begun on July 5, 1973. Both workshop participants and kit recipients were surveyed during the following two weeks.

1. Mail-Out Recipients

All kit recipient calls netted the same results. The recipients had not used the kit. All indicated a desire to do so, but said other obligations were hindering their efforts to look at the kit. When asked if they were willing to be surveyed after using the kit, all said, "yes". Time projected for such use varied from uncertain to next week to six months. This lack of data may, however, provide insight into the ease of kit usage. Experts in the field of evaluation and related fields rated the kit difficult to use and this difficulty may have been too much for the evaluation laymen who received the kits. Change strategy which suggests the need for specific ways to create change should also be examined in this usage problem.
2. Workshop Participants

All workshop participants expressed benefit from their attendance. They appreciated the developer's attempts to meet their individual needs and said they were more interested in evaluation and in continuing to develop their evaluation skills. The development of instruments to meet their needs was very appreciated as was the kit (one commented, "the kit is a much better reference than the notes I usually take at a workshop"). Seventy-five percent of the participants had used the kit since the workshop (this high percentage corresponds to the prediction made on their rating of the kit at the workshop).

An average of 3.25 instruments had been developed per participant and an average of one instrument per participant had been used either during or since the workshop (See Section E). Seventy-five percent of the participants had developed an evaluation design. A pre-post quantitative comparison of instrument and design development revealed that three participants had not constructed instruments before the workshop and none were without after the workshop. A total of eight instruments had been constructed before the workshop and twenty had been constructed after the workshop. Seven workshop participants were without an evaluation design for their instruction before the workshop and only two were without
evaluation designs three to five weeks after the workshop.

Types of instruments developed during and after the workshop varied according to the participant's specific evaluation needs and were categorized in the following fashion; however, all participants were developing more than one type of instrument.

<table>
<thead>
<tr>
<th>Type of Instrument</th>
<th>Number of Instruments Per Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needs Assessment</td>
<td>11</td>
</tr>
<tr>
<td>Process Evaluation</td>
<td>5</td>
</tr>
<tr>
<td>Product Evaluation</td>
<td>4</td>
</tr>
</tbody>
</table>

Needs assessment instruments were those most frequently developed but it was recognized that information obtained would serve as a basis for future process and product instrument development.

Additionally, over half of the participants had used the kit (and workshop) information to help in developing reports on data obtained.

Design components cited by the participants followed the components stressed in the kit. All designs contained a needs assessment, process monitoring and monitoring of outcomes. Information on design components was necessarily general since the survey was by telephone; however, some participants cited specific types of instruments to be utilized in a design (i.e., Gestalt statements or forced choice responses). One participant's
design called for an external evaluator and two participants' designs included plans for the final report. Proper use of terminology, sequence, and instruments in these designs (in light of the lack of design skill on the pre-test) seemed to verify qualitatively other indicators that the workshop and kit combined to provide evaluation skill.

A great increase in self rated achievement was reported with the mean rating improving by more than fifty percent. The workshop apparently made the kit usage easier and with the resultant utilization, achievement continued to increase.

Unanticipated outcomes of the workshop were as follows:

1. Two participants are planning to conduct In-service Training Sessions based upon the workshop and kit this fall.

2. Evaluation "jargon" used effectively in a job interview by one participant.

3. A Title I grant was designed and submitted, using design and instrumentation from the kit and workshop.

Suggestions for additions to the kit and workshop are -- include information on program evaluation and information on graphically reporting data.
E. Rating of Evaluation Products

Two evaluation specialists rated the instruments generated by the workshop participants. Each specialist was given the instruments in a stack. The instruments submitted by participants before the workshop were shuffled with instruments generated after the workshop. Each specialist was asked to separate the instruments into three not necessarily equal groups. The criteria they were asked to use in separating the quality of the instruments were as follows:

a) technical adequacy
b) likely to yield useful information
c) ease to administer
d) ease to analyze
e) ease to interpret

The criteria were very general and each instrument was rated on all criteria simultaneously. Only global impressions could be obtained from this type of assessment but considering the differences among instruments in resources invested, purposes, and types, only global assessments seemed appropriate.

Ten pre and ten post assessment instruments were submitted and rated by the specialists. Few differences seemed to exist between the general quality of the pre-workshop instruments and the post-workshop instruments.
The instruments were rated as follows:

<table>
<thead>
<tr>
<th>Specialist #1</th>
<th>Poor Group</th>
<th>Medium Group</th>
<th>Good Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre Post</td>
<td>Pre Post</td>
<td>Pre Post</td>
</tr>
<tr>
<td>Specialist #1</td>
<td>2 3</td>
<td>4 5</td>
<td>3 3</td>
</tr>
<tr>
<td>Specialist #2</td>
<td>Pre Post</td>
<td>Pre Post</td>
<td>Pre Post</td>
</tr>
<tr>
<td></td>
<td>3 4</td>
<td>4 5</td>
<td>3 1</td>
</tr>
</tbody>
</table>

The specialists agreed in their ratings of eleven instruments; they disagreed by one group on eight instruments; they gave totally opposing ratings on one instrument.
CONCLUSIONS

The Evaluation of Instruction Kit was successful only as a part of the workshop. Utilization of the kit independently has not occurred; however, combined with the workshop, the kit has continued to be a useful resource. Evaluation achievement (self-rated), skills, and knowledge improved as a result of the combined workshop and kit presentation; however, evaluation products (instruments) did not seem to improve.
RECOMMENDATIONS

A longitudinal study should be made after several months to determine if the workshop participants are continuing to use the kit. Unanticipated outcomes, including rippling effects, should be analyzed. A follow-up study of kit recipients should be made to see if they did, in fact, use the kit as indicated in this July survey.
VITA

Helen Louise Blackburn Armstrong

February 28, 1936. . . . . . . Born -

1957 . . . . . . . . . . . . . B.S., Purdue University

1957-1958. . . . . . . . Purdue University, Dean of Men, Secretary

1957-1958. . . . . . . . Willard, Ohio, City Schools, Substitute Teacher and Tutor

1967-1969. . . . . . . . Greenwich, Ohio, South Central High School, Mathematics Teacher

1968 . . . . . . . . . . . . . Ashland College, Education Certification

1969 . . . . . . . . . . . . . Central State University, Education Certification

1969-1970. . . . . . . . Central State University, National Science Foundation Grant

1969-1970. . . . . . . . Xenia, Ohio, West Junior High School, Mathematics Teacher

1970-1971. . . . . . . . Xenia, Ohio, City School System, Evaluation-Dissemination

1971 . . . . . . . . . . . . . Ohio State University, Columbus, Ohio, Evaluation Center, Model Training Project Research Assistant
<table>
<thead>
<tr>
<th>Year Range</th>
<th>Institution and Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-1972</td>
<td>Ohio State University, Columbus, Ohio, Center for Vocational and Technical Education, development of Behavioral Objectives</td>
</tr>
<tr>
<td>1972</td>
<td>M.A., Ohio State University, Columbus, Ohio</td>
</tr>
<tr>
<td>1972-present</td>
<td>Ohio State University, Columbus, Ohio, Nisonger Center, Home Training Consultant Program, Intern Evaluation</td>
</tr>
<tr>
<td>1972-present</td>
<td>Ph.D. Candidate, Ohio State University, Columbus, Ohio</td>
</tr>
</tbody>
</table>
Chapter VI

INTERPRETATIONS BY THE KIT DEVELOPER

Conclusions

The man put the dissertation down and leaned back in his chair. "I read the results by the impact evaluator," he said, "and now I have some questions."

"What are the outstanding ideas in your dissertation? What do you want me to remember about it?"

"I would like for you to remember several things:

The dissertation focuses on the utility of a kit during early development. Although information is collected for revising the kit, revision is not the purpose of the study.

The kit represents an organizational framework, a model. The appropriate test for this type of model is utility.

The dissertation results were encouraging—a majority of workshop participants did use the kit after they were trained.

Parts of the dissertation are intended as a companion to the kit.
Parts of the dissertation are an application of the kit it is validating.

The dissertation is an example of an interim validation study.

The dissertation is evaluative. Therefore, developer and evaluator roles were somewhat divided."

"You seem to have collected much data. Why did you do that if all you really wanted to know was whether the kit was used?"

"I needed to know more than if the kit was used. Only a few people were exposed to the kit. I needed to collect information so that I could describe them. Potential field testers of the kit could then look at the descriptions of participants and decide whether they would like to try to use the kit.

Another reason for collecting so much data was that validity has a wide range of meanings. I wanted to collect information from a wide range of sources.

In addition, since the kit is in early development, I needed to collect a wide range of information to get a stronger impression of whether the final product will have any chance for success; if I only collected information from a narrow range of sources I would have even weaker impressions to base decisions on."
"You once said something about a future study or studies. Why didn't you do both studies at once?"

"It would not have been morally or economically defensible to disseminate the kit to a large number of potential users before information was collected on a small number of people first. A second study is recommended because conclusions on small numbers of people are necessarily limited."

"Okay, so you have much information on a few people. I read through your Results and I read the Impact Evaluation Report carefully. I'm not sure I have a totally clear picture of your results."

The man's wife comes in and says, "Why don't you just call out your results in terms of your objectives. If you don't mind shouting a little I could listen in from the study." (I nod in the affirmative and she leaves.)

"HOW ABOUT IF I JUST LIST EACH OBJECTIVE AND THEN SAY 'POSITIVE' IF I THINK IT WORKED OUT AND 'NEGATIVE' IF IT DIDN'T?" (The wife comes to the door of the study and gives me a look...I lower my voice.)

"well, let me see...remember, you are getting biased judgements..."

Objective #1: Was a descriptive list of resources
developed in evaluation of instruction? **Positive.**

**Objective #2:** Was an evaluation of instruction kit developed? **Positive.**

**Objective #3:** Was an evaluation of instruction workshop developed to train users of the kit? **Positive.**

**Objective #4:** Was the workshop a success? **Positive.**

Specifically:

a. Pre-post measures on skill. **Positive.**

b. Pre-post measures on knowledge. **Positive.**

c. Pre-post measures on attitude toward evaluation. Started positive and ended positive.

d. Pre-post measures on achievement (self-rated). **Positive.**

e. Pre-post quality of evaluation instruments by participants. **Negative (no improvement).**

f. Judgements on the workshop by participants. **Positive.**

g. Participants seeking out evaluation specialists for help after the workshop. **Positive.**

h. Participants with an overall evaluation design after the workshop. **Positive.**

i. Participants using the kit after the workshop. **Positive.**"
Objective #6: Will the kit be used (as it is) without training for potential users? **Negative.**

Objective #7: Did the kit have very much interim validity? **Positive.** Specifically:

- a. Was the kit used? **Positive** (when with training).
- b. Did the kit receive good ratings from experts? **Positive.**
- c. Did the kit receive good ratings from potential users? **Positive.**
- d. Was the utility of the kit demonstrated through its use in this study? **Positive.**

Objective #8: Was a potentially transferable set of procedures for interim validation of an educational product demonstrated? **Positive.**

The wife yells in, "CAN YOU SAY IT ALL IN ONE SENTENCE?"

I give a long pause and then reply, "I think the interim validity of the kit was high—if training goes with the kit; the kit should not be disseminated through the mail; and the workshop training was a success."

"Where do you go from here? Now that the dissertation is completed, what are your intentions regarding the kit?"

"Now that I know that the kit can be useful for some people when it is presented in a workshop, I need to explore the generalizability of the results. Now that I
have a product that brings together what I found as most helpful in evaluation of instruction, the second step seems to be a needs assessment on the national level. A survey needs to be performed to:

a) Discover potential field test sites for the kit;

b) Discover how the kit needs to be presented to increase its likelihood of getting adopted in colleges and universities;

c) Discover if an alternative "answer" should be pursued to assist colleges and universities in evaluation of instruction.

The next step would be to write a grant proposal to finance a full field test of the kit or, if the kit doesn't seem like it will be accepted very broadly, pursue an alternative to the kit."

"If you wrote a grant proposal and it got funded, what information would you be trying to obtain?"

"I would want to find out three things:

First, what kind of people use the kit to perform evaluation of instruction.

Second, do they utilize their results from evaluation? Third, does using their results from evaluation improve educational outputs?"

Before field testing the kit I would want to revise the kit and refine the assessment instruments."
"Do you have any regrets?"

"Well, I certainly got poor results from some decisions. I think I learned much about what I can do though--such as what time of day I can expect to give a good workshop and when I can expect to give a poor one. I learned how my perceptions of my work differs from perceptions of others in different contexts. So, no--I can't say I have regrets. I don't think I would want to do it all over again, however. I feel very fortunate that the workshop went well and that some colleagues were able to find the kit of use."

"I would like to see a list of specific recommendations you would make as a result of your study."
Recommendations

"The main recommendation is to conduct more workshops. Reasons include the following:

1. Test revisions of kit, especially kit instructions.
2. Test workshop revisions designed to get rid of jargon problems, logistics problems, transition problems, and communication problems on participant activities.
3. Aggregate data to allow greater generalizability of results.
4. Aggregate data to establish predictive validity on achievement instrument.
5. Control for possible Hawthorne Effects.
6. Control for intensity of possible "halo" effects.
7. Reassess content validity.
8. Get more qualitative data on impact of workshop - for instance, quality of participant instruments, where most typical flaws are, quality of evaluation designs by participants.
9. Find out more about which parts of the kit tend to get used most.
10. Begin dissemination of the kit.
11. Discover more about specific needs of potential audience.
12. Gain some indications on the effects of the kit on improving student benefits from instruction.
A second major recommendation is to refine assessment devices used in the study. Activities include the following:

1. Generate alternate forms on knowledge and skill instruments in order to assess learnings with some generalizability.

2. Establish predictive validity of item four of Rating Evaluation of Instruction instrument so that follow-up may eventually be unnecessary.

3. Establish construct and predictive validity on an instrument to assess attitudes toward evaluation of instruction in order to select future participants for the workshops.

4. Refine criteria for judging responses to skill and knowledge instruments, correlate results with achievement instrument, and then use skill and knowledge instruments solely for training purposes.

5. Establish inter-judge reliability on expert judges used — for example, raters on technical adequacy of the kit. Criteria for judgements need to be clarified and separated before reliability results will be very meaningful.

A third recommendation is to refine the procedures used in conducting this type of study. Specific areas that need attention include the following:

1. Revise flow chart of activities to improve
procedures for data collection and selection of participants.

2. Generate a more useable PERT chart so that support personnel, such as secretaries, can follow it independently.

3. Generate and pilot test written instructions for all support personnel.

4. Perform follow-up by mail and link responses to strong incentives, such as grades, remission of registration charge, etc.

5. Specify the tasks of the summative evaluator very concretely: specify which instruments are to be involved, which comparisons are to be made, which data analysis procedures are to be used, and what context and standards are to be used for making judgements. A contract should be written which includes a penalty clause for extension of the deadline for the report. The contract should also contain other relevant data contained in the Evaluation Contracting manual by Stephen Klein.

A fourth recommendation, one which relates to the development effort rather than the validation focus, involves revision activities. Both the kit and the workshop should be revised. The kit might be revised in the following ways:
1. Give more tutorial tests on instructions to use the kit.
2. Follow-up tutorial tests and collect data on use of kit without workshop.
3. Add guidelines on item construction into the kit.
4. Add an attractive synopsis of results on the validation study (dissertation).
5. Revise resources section to include parts of review of literature in dissertation.
6. Provide a list of definitions in the kit.
7. Provide a list of the kit developer's assumptions.
8. Delete advice on statistical analysis and control groups - give resources instead.
9. Provide a short (paragraph) statement on how each instrument might typically be used.
10. Give both a generalizable form and an example for each instrument and clearly distinguish between the two.
11. Provide a case study example on how to use the kit.
12. Generate narrative on validation, needs assessment, and criterion referenced testing.
The workshop might be revised in the following ways:

1. Provide for interactions among participants early in the workshop and in a way that is helpful for them in using each other's experience and training.

2. Provide a list of definitions of evaluation jargon early in the workshop; go over the definitions orally.

3. Write out directions for transitions among workshop activities and pilot test instructions.

4. Don't lecture before 11:00 A.M. and preferably not before 1:00 P.M. (for me).

5. If possible, "seed" each work group with at least one enthusiastic participant.

6. Narrow the scope of workshop objectives.

7. Make the objectives and needs assessment items totally isomorphic; delete material not directly related to objectives.

8. Explain the theoretical basis of the kit and how to use the kit early in the workshop.

9. Find ways to make the workshop more enjoyable and not so constantly "heavy."

In addition to the above revisions, some attention should be given to components of the workshop that should be continued:

Continue to use good consultants and support staff
during the workshop.

Continue to allow for basic changes in content or format to meet participant needs.

Continue to systematically monitor the progress of the workshop.

Continue to concentrate on participant productivity."

"Those are fine recommendations. Maybe I will see you again sometime," the man said, as he turned to leave.

"Wait! Who the heck are you?" I was too late. The man was disappearing way up, up in the air and crawled into what for all the world looked like an Ivory Tower.


Baker, Eva L. Preparing Instructional Materials for Educational Developers. Los Angeles, California: Graduate School of Education, University of California (under development).


Collet, Leverne S. *Fehr-Practicum.* Ann Arbor, Michigan: 4215A School of Education Building, University of Michigan (under development).


Cronback, Lee J. "Course Improvement Through Evaluation." *Teachers College Record,* Vol. 64, 1963, 672-83.


Jung, Charles; Pino, Rene; and Emory, Ruth. Research Utilization and Problem Solving. Portland, Oregon: Northwest Regional Educational Laboratory, 1970.


Milczarek, Gary J. *Instructional Development Formative Evaluation Design for EPSC*. Columbus, Ohio: The Ohio State University.


Morgan, James M. *A Seminar and Training Program in Needs Assessment and Goal Development*. Cincinnati, Ohio: Division of Program Research and Design, Evaluation Center (under development).


Stufflebeam, Daniel L. "An Introduction to the Theories, Functions, and Methods of Educational Evaluation." Columbus, Ohio: The Ohio State University, College of Education Evaluation Center, 1971 (Typewritten Draft).


APPENDICES:

Documents to Demonstrate Use of the Evaluation of Instruction Kit and Interim Validation Procedures.
Appendix A ......................................................... 190
  1. Before the Workshop (letters, instruments, flow chart)*
  2. Before the Kit Mail-Out (letters, instruments, flow chart)**
Appendix B ......................................................... 202
  1. Beginning the Workshop (agenda and instruments)
  2. Before Opening the Mailed Kit (letter and instruments)
Appendix C .......................................................... 220
  1. Throughout the Workshop (instruments)
Appendix D .......................................................... 225
  1. End of Workshop (instruments and reports)
Appendix E .......................................................... 248
  1. After the Workshop (follow-up letter and instruments)
  2. After Receiving the Mailed Kit (follow-up letter and instruments)
Appendix F .......................................................... 266
  1. Instrument for Assessment of Workshop Instruments
  2. Instrument for Assessment of Kit Instruments
  3. Letter and Instrument for Assessment of Kit by Experts

*1= Information on workshop participants
**2= Information on mail-out recipients
APPENDIX A:
1. Before the Workshop
TO Faculty and Staff:

In cooperation with the State Division of Mental Retardation and Developmental Disabilities, the Evaluation Unit of Nisonger Center is offering a workshop on evaluation. This workshop is designed to assist developers of instruction, instructors, and administrators in continuing and higher education with evaluation needs. At the end of the workshop participants should be able to:

1. Describe an overview of evaluation relevant to their particular needs;
2. Develop some of their own evaluation instruments;
3. Generate much of their own evaluation design; and
4. Use an evaluation specialist efficiently as a resource.

The workshop has been designed to validate an evaluation kit and is therefore experimental in nature. No registration fee will be requested for this first workshop. However, since the workshop is experimental in nature, only a limited number of participants will be accepted. Applications will not be accepted after May 30, 1971.

If you or someone on your staff would like to attend the workshop, please fill out the enclosed registration application form and background information sheet. The background information sheet is designed to: (1) help individualize the workshop to the particular participants attending and (2) help the developers insure that a full range of backgrounds are represented.

For additional information call or write to:

Jerry Adams
Evaluation Unit, Rm. 439
Nisonger Center
1580 Cannon Drive
Columbus, Ohio 43210

Phone: 614-422-9780
Registration Application

The three-day Evaluation Workshop will be held in Columbus, Ohio (Nisoner Center). Place a check beside the workshop dates which are open for you. If more than one date is open, also rank the dates in order of preference. (*=most preferred)

____ a) June 11, 12 and 13*
____ b) June 18, 19 and 20*
____ c) July 9, 10 and 11*
____ d) July 16, 17 and 18*

Name: ____________________________________________

Address: __________________________________________

Organization: _______________________________________

Position: __________________________________________

Return registration application and background information to:

Jerry Adams
Pm, 439 Nisoner Center
1580 Cannon Dr.
Columbus, Ohio 43210

You will be notified soon after May 23 concerning your acceptance.

Note: Return your Background Information with your Registration Application.

* Day #1 8:30-5:00
   Day #2 8:00-5:00
   Day #3 8:30-4:00
BACKGROUND INFORMATION

1. What will you be doing in June, 1973? (Check all that apply)

   ■ Teaching higher education course(s)
   ■ Teaching continuing education course(s)
   ■ Developing instructional materials
   ■ Involved in project or program evaluation
   ■ None of the above

Circle the number that best describes your level of competency for each of the following items. The responses are divided into five categories.

1 = No competency. Respondent has no training or experience base from which to work.

2 = Minimum Competency. Respondent has had some training and experience in this area, but does not feel confident enough to perform the task without further training.

3 = Moderate Competency. Respondent has fairly thorough training and experience in this area and could perform the task moderately well.

4 = High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.

5 = Very High Competency. Respondent is completely confident in performing this task and has done creative work in this area.

2. Develop instrument to logically analyze instructional materials

3. Develop needs assessment instrument

4. Develop criterion referenced test

5. Develop instrument to monitor the instructional process

6. Develop instrument to collect reactions from participants

7. Develop an instrument to assess impact, including unintended outcomes

8. Develop an instrument to assess long-term effects

9. Develop evaluation design

10. Find needed resources for data analysis

11. Validate an instrument

12. Interpret evaluation results

13. Report evaluation results

14. Level of competency to do educational evaluation from previous job experience

15. Level of competency to do educational evaluation from previous training
Selection Procedures for Workshop Participants

START

June 11-13 OK?
  yes
  involved in evaluation of instruction?
  yes
    in June?
    yes
    invite to workshop; ask to bring materials

  no
  sorry, send letter

  no
  sorry, send letter

  need hotel reservations?
  yes
  make reservations

  no
  send workshop confirmation letter

  send confirmation letter and map
THE OHIO STATE UNIVERSITY

Jerry Adams
Evaluation Unit Room #439
Mixonger Center
1580 Cannon Drive
Columbus, Ohio 43210

Acceptance: Workshop

Dear (Columbus' Residents)

This is to confirm your acceptance in the evaluation workshop on June 11th, 12th and 13th (first session starts at 8:30).

Please Note:

1. Bring whatever you have developed for evaluation of your instruction. Much time can be saved by building on what you already have.

2. Bring instructional materials (notes, etc.) you now use that you would like to work with during the workshop.

3. The workshop format will mostly involve work groups. Lecture will be minimal. Work sessions will build on work of former sessions. The workshop is designed to help you build a product (an evaluation design and instruments) for your immediate use after the workshop. Partial attendance is therefore discouraged.

Sincerely,

Jerry Adams

JA/cm
Jerry Adams  
Evaluation Unit Room 3439  
Memonger Center  
1580 Cannon Drive  
Columbus, Ohio  43210

Acceptance: Workshop

Dear (out of town residents)  

This is to confirm your acceptance in the evaluation workshop on June 11th, 12th and 13th (first session starts at 8:30).  

Reservations have been made for you at ________________.

Please Note:

1. Bring whatever you have developed for evaluation of your instruction. Much time can be saved by building on what you already have.

2. Bring instructional materials (notes, etc.) you now use that you would like to work with during the workshop.

3. The workshop format will mostly involve work groups. Lecture will be minimal. Work sessions will build on work of former sessions. The workshop is designed to help you build a product (an evaluation design and instruments) for your immediate use after the workshop. Partial attendance is therefore discouraged.

Sincerely,

Jerry Adams

JA/cm
APPENDIX A:
2. Before the Kit Mail-out
TO Faculty and Staff:

In cooperation with the State Division of Mental Retardation and Developmental Disabilities, the Evaluation Unit of Nisonger Center is offering a workshop on evaluation. This workshop is designed to assist developers of instruction, instructors, and administrators in continuing and higher education with evaluation needs. At the end of the workshop participants should be able to:

1. Describe an overview of evaluation relevant to their particular needs;
2. Develop many of their own evaluation instruments;
3. Generate much of their own evaluation design; and
4. Use an evaluation specialist efficiently as a resource.

The workshop has been designed to validate an evaluation kit and is therefore experimental in nature. No registration fee will be requested for this first workshop. However, since the workshop is experimental in nature, only a limited number of participants will be accepted. Applications will not be accepted after May 30, 1973.

If you or someone on your staff would like to attend the workshop, please fill out the enclosed registration application form and background information sheet. The background information sheet is designed to:
(1) help individualize the workshop to the particular participants attending and (2) help the developers insure that a full range of backgrounds are represented.

For additional information call or write to:

Jerry Adams
Evaluation Unit, Rm. 439
Nisonger Center
1580 Cannon Drive
Columbus, Ohio 43210

Phone: 614-422-9780
Registration Application

The three-day Evaluation Workshop will be held in Columbus, Ohio (Hisonger Center). Place a check beside the workshop dates which are open for you. If more than one date is open, also rank the dates in order of preference. (1=most preferred)

_____ a) June 11, 12 and 13*
_____ b) June 18, 19 and 20*
_____ c) July 9, 10 and 11*
_____ d) July 16, 17 and 18*

Name: __________________________________________

Address: __________________________________________

________________________________________________________________________

Organization: __________________________________________

Position: ____________________________________________

Return registration application and background information to:

Jerry Adams
Rm. 639 Hisonger Center
1580 Cannon Dr.
Columbus, Ohio 43210

You will be notified soon after May 30 concerning your acceptance.

Note: Return your Background Information with your Registration Application.

* Day #1 8:30 - 5:00
   Day #2 8:00 - 5:00
   Day #3 8:30 - 4:00
BACKGROUND INFORMATION

1. What will you be doing in June, 1973? (Check all that apply)
   ______ Teaching higher education course(s)
   ______ Teaching continuing education course(s)
   ______ Developing instructional materials
   ______ Involved in project or program evaluation
   ______ None of the above

Circle the number that best describes your level of competency for each of the following items. The responses are divided into five categories.

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No competency. Respondent has no training or experience base from which to work.</td>
</tr>
<tr>
<td>2</td>
<td>Minimum Competency. Respondent has had some training and experience in this area, but does not feel confident enough to perform the task without further training.</td>
</tr>
<tr>
<td>3</td>
<td>Moderate Competency. Respondent has fairly thorough training and experience in this area and could perform the task moderately well.</td>
</tr>
<tr>
<td>4</td>
<td>High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.</td>
</tr>
<tr>
<td>5</td>
<td>Very High Competency. Respondent is completely confident in performing this task and has done creative work in this area.</td>
</tr>
</tbody>
</table>

2. Develop instrument to logically analyze instructional materials......1 2 3 4 5
3. Develop needs assessment instrument.........................................1 2 3 4 5
4. Develop criterion referenced test..............................................1 2 3 4 5
5. Develop instrument to monitor the instructional process..............1 2 3 4 5
6. Develop instrument to collect reactions from participants............1 2 3 4 5
7. Develop an instrument to assess impact, including unintended outcomes........................................1 2 3 4 5
8. Develop an instrument to assess long-term effects........................1 2 3 4 5
9. Develop evaluation design..........................................................1 2 3 4 5
10. Find needed resources for data analysis....................................1 2 3 4 5
11. Validate an instrument.............................................................1 2 3 4 5
12. Interpret evaluation results.....................................................1 2 3 4 5
13. Report evaluation results........................................................1 2 3 4 5
14. Level of competency to do educational evaluation from previous job experiences........................................1 2 3 4 5
15. Level of competency to do educational evaluation from previous training....................................................1 2 3 4 5
Selection Procedures for Mail-Out Recipients

START

June 11-13 OK?

- no
  - involved in evaluation of instruction?
    - no → sorry, send letter
    - yes → in June?
      - no → sorry, send letter
      - yes → will send kit
        → send pre-test and letter
        → send kit
APPENDIX B:
1. Beginning the Workshop
WORKSHOP ON EVALUATION OF INSTRUCTION

Monday, June 11

Agenda

8:30 - 9:00  Registration/Coffee
9:00 - 9:45  Welcome: Bill Gibson, Director of Nisonger
             Introduction and Orientation: Jerry Adams
             (Evaluation Consultants): Gary Milczarek
             Unhai Ahn
9:45 - 10:45 Individual Work: Participant Assessment
             Examine Kit and Resources
10:45 - 11:15 Lecture: Needs Assessment
11:15 - 11:30 Large Group: Brainstorm on skills most needed for
             evaluation of instruction/Prioritize
11:30 - 12:45 Lunch
1:00 - 2:00  Individual Work: Develop Needs Assessment Instrument
2:00 - 2:30  Small Groups (triads): Critique Instruments
2:30 - 3:00  Individual Work: Revise Instruments -- Hand in.
3:00 - 3:30  Small Groups: Analysis of Materials
3:30 - 3:50  Gestalt Statements/ Break
3:50 - 4:45  Presentation: Evaluation Kit
4:45 - 4:55  Summary of the Day
4:55 - 5:00  Complete Topic Feedback Instrument
Tuesday, June 12

8:00 - 8:30  Summary -- Presentation -- Item Construction
by Jerry Adams

8:30 - 9:00  Instrument Development

9:00 - 10:00  Critique

10:00 - 10:15  Break

10:15 - 10:30  Gestalt Statements

10:30 - 11:00  Presentation: Types of Instruments, By Jerry Adams

11:00 - 11:30  Discussion

11:30 - 1:00  Lunch

1:00 - 1:45  Instrument Development

1:45 - 2:30  Critique, Rewrite, Hand In

2:30 - 3:00  Presentation -- Validity, By Jerry Adams and
Bill Loadman

3:00 - 3:15  Break

3:15 - 3:45  Instrument Development

3:45 - 4:15  Critique, Rewrite, Hand In

4:15 - 4:45  Presentation: Evaluation Issue by Jerry Adams

4:45 - 5:00  Goals for Session Instruments

Wednesday, June 13

8:30 - 9:00  Review of Activities
Presentation -- Debriefing Instrument

9:00 - 9:30  Develop Debriefing Form (using kit)

9:30 - 10:00  Debrief and Critique

10:00 - 10:15  Break

10:15 - 10:45  Open Forum

10:45 - 11:30  Evaluation Report

11:30 - 11:35  Group Effectiveness Instrument

11:35 - 1:00  Lunch

1:00 - 1:30  Presentation: Evaluation Design by Jerry Adams

1:30 - 2:30  Workshop Summary by Jerry Adams

3:00 - 4:00  Assessment

4:00 - 5:00  Debriefing
INSTRUMENTS
Name

Code No.

Workshop: Evaluation of Instruction

Background Information

Please circle the number in front of responses as they apply to you:

1. Sex: 1. Male
     2. Female

2. Employer:
     1. State Department
     2. College or University
     3. State School
     4. Research and Development Lab
     5. Other (specify)

3. Principal role in present position:
     1. Administrator
     2. Evaluator
     3. Instructor
     4. Developer
     5. Other (specify)

4. Highest degree attained
     1. None
     2. Associate
     3. Bachelor's
     4. Master's
     5. Doctoral

Check

Yes No

5. Do you have some instruments already developed for evaluation of instruction? If yes, please list them; if no, go to "7."

   1.  
   2.  
   3.  
   4.  
   5.  
   6.  

6. Please place a check by each of those (above) that you have used.
   Did you bring the instruments that you developed?

   ___ ___

7. Do you have an overall design for evaluating instruction? If yes, please outline the components below:
Please assist the planning of the Evaluation of Instruction Workshop by circling the number that best describes the level of competency you intend to achieve within the next two weeks for each of the following items.

The responses are divided into five categories:

1=No Competency. Respondent has had no training and experience base from which to work.
2=Minimum Competency. Respondent has had some training or experience in this area, but does not feel confident enough to perform the task without further training.
3=Moderate Competency. Respondent has fairly thorough training or experience in this area and could perform the task moderately well.
4=High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.
5=Very High Competency. Respondent is completely confident in performing this task and has done creative work in this area.

<table>
<thead>
<tr>
<th>Task</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop instrument to logically analyze instructional materials</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop needs assessment instrument</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop criterion referenced test</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop instrument to monitor the instructional process</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop instrument to collect student reactions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop instrument to assess the impact of instruction, including unintended outcomes</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop instrument to assess the long-term effects of instruction</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Develop evaluation design for instruction</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Find needed resources for data analysis</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Validate an instrument</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Interpret evaluation results</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Report evaluation results</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Level of competency to do educational evaluation from previous job experiences</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Level of competency to do educational evaluation from previous training</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Entry Knowledge

1. List the first three steps you would take to generate a needs assessment instrument:
   1. 
   2. 
   3. 

2. List the first three steps you would take to generate a criterion referenced test:
   1. 
   2. 
   3. 

3. List two kinds of instruments to monitor the instructional process:
   1. 
   2. 

4. List two methods for collecting reactions from participants:
   1. 
   2. 

5. List two resources you might use to assist you with data analysis:
   1. 
   2. 

6. List the main three topics or components you would include in an evaluation report:
   1. 
   2. 
   3. 

7. List types of instrument validity:
2-C

1. Generate three items for use in a needs assessment instrument:
   Directions: ____________________________________________________________
   ____________________________________________________________________
   1. ________________________________________________________________
   2. ________________________________________________________________
   3. ________________________________________________________________

2. Generate three items for use in a criterion-referenced instrument:
   Directions: __________________________________________________________
   ____________________________________________________________________
   1. ________________________________________________________________
   2. ________________________________________________________________
   3. ________________________________________________________________

3. Generate an instrument format to monitor the instructional process.
   (Use back of page)

4. Generate three items for use on an instrument to logically analyze
   instructional materials:
   Directions: __________________________________________________________
   ____________________________________________________________________
   1. ________________________________________________________________
   2. ________________________________________________________________
   3. ________________________________________________________________

5. Generate three items to use in an instrument to collect reactions from participants:
   Directions: __________________________________________________________
   ____________________________________________________________________
   1. ________________________________________________________________
   2. ________________________________________________________________
   3. ________________________________________________________________

6. Generate an item for use in an instrument to assess impact of instruction, including unintended outcomes:
7. Generate three items for use in an instrument to assess long-term effects:

Directions:_________________________________________________________

1._______________________________________________________________

2._______________________________________________________________

3._______________________________________________________________
Directions: This opinionnaire attempts to assess attitudes of participants toward educational evaluation. Please indicate your agreement or disagreement with each statement by checking ( ) the answer that best describes how you personally feel, regardless of whether other people may agree or disagree with you. Although many of the questions may appear similar, please judge each one on an individual basis. Since we need to know your attitude, please answer each question frankly and honestly. There are no correct responses. There is no time limit, but respond as quickly as you can, and do not leave out any of the statements.

1. I see and feel deeply a need for increased evaluation information for decision-making in my field.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

2. Educational evaluation usually results in arbitrary judgements about the educative process.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

3. There is currently too much concern with evaluation in education.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

4. Educators do not receive sufficient training in evaluation.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

5. Educative processes are too complex to be evaluated objectively.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

6. Evaluation is an attempt to reduce education to a mechanistic process.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral

7. Educational evaluation usually results in improvements in educational practice.

( ) ( ) ( ) ( ) ( )
Strongly Disagree Undecided or Agree Strongly Agree
Disagree Neutral
8. Evaluation should aid an educator in revising his goals even while the program is in progress.

( ) Strongly Disagree
( ) Undecided or Neutral
( ) Agree
( ) Strongly Agree

9. Evaluation interferes with the running of schools more than helps.

( ) Strongly Disagree
( ) Undecided or Neutral
( ) Agree
( ) Strongly Agree

10. Intuition and general knowledge of practitioners are more valuable than formal evaluative procedures in making decisions in education.

( ) Strongly Disagree
( ) Undecided or Neutral
( ) Agree
( ) Strongly Agree

11. Using educational evaluation in my work does not appeal to me.

( ) Strongly Disagree
( ) Undecided or Neutral
( ) Agree
( ) Strongly Agree

12. Money spent on evaluation contributes more to the improvement of education than any other expenditure.

( ) Strongly Disagree
( ) Undecided or Neutral
( ) Agree
( ) Strongly Agree

*Averill, M., Werther, B. Evaluation of AERA Travelling Institute No. 2.
Criteria of Participants

1. Which criteria will you use in assessing your performance? Please rank the criteria (1 = most important)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>____</td>
<td>a. quality of instruments produced</td>
</tr>
<tr>
<td>____</td>
<td>b. self-rating of achievement at end of workshop</td>
</tr>
<tr>
<td>____</td>
<td>c. success in generating an evaluation design</td>
</tr>
<tr>
<td>____</td>
<td>d. skill in utilizing an evaluation specialist</td>
</tr>
<tr>
<td>____</td>
<td>e. success in utilizing own instruments</td>
</tr>
<tr>
<td>____</td>
<td>f. success in applying an evaluation design</td>
</tr>
<tr>
<td>____</td>
<td>g. other (specify)</td>
</tr>
<tr>
<td>____</td>
<td>h. other (specify)</td>
</tr>
</tbody>
</table>

2. Which format of instruction (if any) do you generally prefer? (Please rank the formats (1 = most preferred).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>____</td>
<td>a. reading</td>
</tr>
<tr>
<td>____</td>
<td>b. small group task force</td>
</tr>
<tr>
<td>____</td>
<td>c. large group discussion</td>
</tr>
<tr>
<td>____</td>
<td>d. individual tasks</td>
</tr>
<tr>
<td>____</td>
<td>e. lecture with media</td>
</tr>
<tr>
<td>____</td>
<td>f. combinations of the above</td>
</tr>
<tr>
<td>____</td>
<td>g. other (specify)</td>
</tr>
<tr>
<td>____</td>
<td>h. other (specify)</td>
</tr>
</tbody>
</table>

3. Which criteria do you think should be used to assess the worth of a course of this type? Please rank the criteria (1 = most important).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>____</td>
<td>relevance to your present position</td>
</tr>
<tr>
<td>____</td>
<td>anticipated relevance to your future position(s)</td>
</tr>
<tr>
<td>____</td>
<td>value compared to other instruction in the same subject</td>
</tr>
<tr>
<td>____</td>
<td>effect of instruction on your professional activity</td>
</tr>
<tr>
<td>____</td>
<td>degree to which instructor met his instructional objectives</td>
</tr>
</tbody>
</table>
Criteria of Participants (continued)

___ amount of inquiry instruction stimulated

___ transferability of knowledge or skills to other areas of your life

___ other ____________________________

___ other ____________________________

___ other ____________________________
APPENDIX B:
2. Before Opening the Mailed Kit
THE OHIO STATE UNIVERSITY

Jerry Adams
Evaluation Unit Room #439
McGraw Center
1580 Cannon Drive
Columbus, Ohio 43210

Acceptance Kit

Dear

This is to confirm that you are unable to attend the evaluation workshop on the date it is being held. We would, however, like to offer some service to you. We are sending you the evaluation kit. We will contact you in two to three weeks to find out how useful the kit has been. For those who have attempted to utilize some of the kit materials or ideas we will offer consultation services if requested. It is our understanding that you have need for evaluation of instruction activities during June (We assume that you will begin using the kit within a few days after receiving it). We wish you well in your efforts.

We would like some assistance from you. Enclosed are several evaluation instruments. Please fill them out and return them to Jerry Adams before opening the kit. The results from the enclosed instruments will help us be clearer about the amount of assistance that the kit gave to you during June.

Sincerely,

Jerry Adams

JA/cm
Name__________________
Code No.________________
Workshop: Evaluation of Instruction

Background Information

Please circle the number in front of responses as they apply to you:

1. Sex: 1. Male
   2. Female

2. Employer:
   1. State Department
   2. College or University
   3. State School
   4. Research and Development Lab
   5. Other (specify)____________________

3. Principal role in present position:
   1. Administrator
   2. Evaluator
   3. Instructor
   4. Developer
   5. Other (specify)____________________

4. Highest degree attained
   1. None
   2. Associate
   3. Bachelor's
   4. Master's
   5. Doctoral

   Check
   Yes  No

5. Do you have some instruments already developed for evaluation of instruction? If yes, please list them; if no, go to 27.
   1. ______________________________
   2. ______________________________
   4. ______________________________
   5. ______________________________
   6. ______________________________

6. Please place a check by each of those (above) that you have used.

   Check
   Yes  No

7. Do you have an overall design for evaluating instruction? If yes, please outline the components below:
Entry Knowledge

1. List the first three steps you would take to generate a needs assessment instrument:
   1. 
   2. 
   3. 

2. List the first three steps you would take to generate a criterion referenced test:
   1. 
   2. 
   3. 

3. List two kinds of instruments to monitor the instructional process:
   1. 
   2. 

4. List two methods for collecting reactions from participants:
   1. 
   2. 

5. List two resources you might use to assist you with data analysis:
   1. 
   2. 

6. List the main three topics or concepts you would include in an evaluation report:
   1. 
   2. 
   3. 

7. List types of instrument validity:
Entry Skill

1. Generate three items for use in a needs assessment instrument:
   Directions:__________________________________________________________
   _________________________________________________________________
   1._______________________________________________________________
   2._______________________________________________________________
   3._______________________________________________________________

2. Generate three items for use in a criterion-referenced instrument:
   Directions:________________________________________________________
   _________________________________________________________________
   1._______________________________________________________________
   2._______________________________________________________________
   3._______________________________________________________________

3. Generate an instrument format to monitor the instructional process.
   (use back of page)

4. Generate three items for use on an instrument to logically analyze
   instructional materials:
   Directions:________________________________________________________
   _________________________________________________________________
   1._______________________________________________________________
   2._______________________________________________________________
   3._______________________________________________________________

5. Generate three items to use in an instrument to collect reactions
   from participants:
   Directions:________________________________________________________
   _________________________________________________________________
   1._______________________________________________________________
   2._______________________________________________________________
   3._______________________________________________________________

6. Generate an item for use in an instrument to assess impact of in-
   struction, including unintended outcomes:
Code no. Entry Attitude

Opinion Survey* (Example)

Directions: This opinionnaire attempts to assess attitudes of participants toward educational evaluation. Please indicate your agreement or disagreement with each statement by checking ( ) the answer that best describes how you personally feel, regardless of whether other people may agree or disagree with you. Although many of the questions may appear similar, please judge each one on an individual basis. Since we need to know your attitude, please answer each question frankly and honestly. There are no correct responses. There is no time limit, but respond as quickly as you can, and do not leave out any of the statements.

1. I see and feel deeply a need for increased evaluation information for decision-making in my field.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

2. Educational evaluation usually results in arbitrary judgements about the educative process.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

3. There is currently too much concern with evaluation in education.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

4. Educators do not receive sufficient training in evaluation.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

5. Educative processes are too complex to be evaluated objectively.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

6. Evaluation is an attempt to reduce education to a mechanistic process.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral

7. Educational evaluation usually results in improvements in educational practice.

( ) Strongly Disagree ( ) Undecided or ( ) Agree ( ) Strongly Agree

Disagree Neutral
8. Evaluation should aid an educator in revising his goals even while the program is in progress.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

9. Evaluation interferes with the running of schools more than helps.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

10. Intuition and general knowledge of practitioners are more valuable than formal evaluative procedures in making decisions in education.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

11. Using educational evaluation in my work does not appeal to me.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

12. Money spent on evaluation contributes more to the improvement of education than any other expenditure.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

APPENDIX C: Throughout the Workshop
1. Instruments
GESTALT STATEMENTS
(anonymous)

Directions: Complete the following statements. Give your immediate reaction to any aspect of the session.

I appreciate

I resent

I want

I learned

I plan

I wish

I feel
1. What topic (if any) presented during this session would you like to see expanded?

2. What topic (if any) presented during this session would you like to see deleted?

3. New terms presented that need further definition or explanation:

Please circle the letter representing the appropriate answer for you on each of the following questions.

4. The most helpful format for me for covering this topic would be:
   a. lecture and discussion
   b. panel presentation and interaction with panel
   c. role play and discussion
   d. whole group discussion
   e. combination of small group and large group discussion

5. This session was very relevant to me in my present position
   a. strongly agree
   b. agree
   c. neutral
   d. disagree
   e. strongly disagree

6. The theoretical framework presented in this session was very useful
   a. strongly agree
   b. agree
   c. neutral
   d. disagree
   e. strongly disagree

7. The specific suggestions and advice presented in this session met my needs
   a. strongly agree
   b. agree
   c. neutral
   d. disagree
   e. strongly disagree

8. The most important issues for me (relative to the topic of the session) were addressed
   a. strongly agree
   b. agree
   c. neutral
   d. disagree
   e. strongly disagree

9. The length of the session was
   a. much too long
   b. too long
   c. just right
   d. too short
   e. much too short

10. I would be interested in participating in a short interview (between some participants and the presenter of the session) in order to work on the development of this session. __Yes  __No
Goals for Participants

1. Goals for participants:
   - Goal 1: Improve skill in needs assessment
   - Goal 2: Improve skill in generating an instrument for assessing skill, knowledge, affect, achievement, or unintended outcomes.
   - Goal 3: Improve skill in generating an instrument to monitor the instructional process.

2. Of the goals listed above, which seem to have been attained or partially attained by you? (You may circle more than one)
   a. 1
   b. 2
   c. 3
   d. none

3. If your answer to the above question was not (d), attempt to assess the degree to which the goal(s) was met. Place a check.

DEGREE

<table>
<thead>
<tr>
<th></th>
<th>Great</th>
<th>Partial</th>
<th>Minimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal 3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Group Effectiveness Scale

1. Overall Effectiveness as a Group Member

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
</table>

2. To what extent does he/she influence the group? (Circle and initial for each member)

- No Influence
- Some Influence
- Great Deal of Influence

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

3. To what extent did the team work at discovering and using the unique backgrounds and role of individuals?

- Did Not Work
- Discovered & Used at All
- Discovered & Used Part of Time
- Discovered & Used Effectively

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

4. How productive was the work of the group?

- Completely Unproductive
- Half as Productive as They Could Have Been
- Very Productive as Much as Possible

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

5. Was there active testing and building on each other's ideas?

- Each Tried to Get About 50-
- His Own Agenda 50
- Across Desire to Complement & Help Each Other Very Evident, Built on Each Other's Ideas.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

APPENDIX D:
1. End of Workshop
   a. Instruments
Personally administered by presenter to entire group of workshop participants

DEBRIEFING FORM

Instructor: "I would like to explore with you your responses on the instruments I have used in this workshop. After looking through the responses of the participants, I would like for you to help me solve the following problems:"

1. ______________________________________
   ______________________________________
   Student reaction and recommendations:_______
   ______________________________________
   ______________________________________
   ______________________________________

2. ______________________________________
   ______________________________________
   Student reactions and recommendations:_______
   ______________________________________
   ______________________________________
   ______________________________________

Exit Knowledge

1. List the first three steps you would take to generate a needs assessment instrument:
   1. ________________________________
   2. ________________________________
   3. ________________________________

2. List the first three steps you would take to generate a criterion referenced test:
   1. ________________________________
   2. ________________________________
   3. ________________________________

3. List two kinds of instruments to monitor the instructional process:
   1. ________________________________
   2. ________________________________

4. List two methods for collecting reactions from participants:
   1. ________________________________
   2. ________________________________

5. List two resources you might use to assist you with data analysis:
   1. ________________________________
   2. ________________________________

6. List the main three topics or components you would include in an evaluation report:
   1. ________________________________
   2. ________________________________
   3. ________________________________

7. List types of instrument validity:
EXIT ACHIEVEMENT (SELF-REPORT)

Please assist the planning of the next Evaluation of Instruction Workshop.

Circle the number that best describes your level of competency for each of the following items. The responses are divided into five categories:
1=No Competency. Respondent has had no training and experience base from which to work.
2=Minimum Competency. Respondent has had some training or experience in this area.
3=Moderate Competency. Respondent has fairly thorough training or experience in this area and could perform the task moderately well.
4=High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.
5=Very High Competency. Respondent is completely confident in performing this task and has done creative work in this area.

1. Develop instrument to logically analyze instructional materials

2. Develop needs assessment instrument

3. Develop criterion referenced test

4. Develop instrument to monitor the instructional process

5. Develop instrument to collect student reactions

6. Develop instrument to assess the impact of instruction, including unintended outcomes

7. Develop instrument to assess the long-term effects of instruction

8. Develop evaluation design for instruction

9. Find needed resources for data analysis

10. Validate an instrument

11. Interpret evaluation results

12. Report evaluation results
Exit Skill

1. Generate three items for use in a needs assessment instrument:
   Directions:________________________________________________________________________
   1.______________________________________________________________________________
   2.______________________________________________________________________________
   3.______________________________________________________________________________

2. Generate three items for use in a criterion-referenced instrument:
   Directions:________________________________________________________________________
   1.______________________________________________________________________________
   2.______________________________________________________________________________
   3.______________________________________________________________________________

3. Generate an instrument format to monitor the instructional process. (Use back of page)

4. Generate three items for use on an instrument to logically analyze instructional materials:
   Directions:________________________________________________________________________
   1.______________________________________________________________________________
   2.______________________________________________________________________________
   3.______________________________________________________________________________

5. Generate three items to use in an instrument to collect reactions from participants:
   Directions:________________________________________________________________________
   1.______________________________________________________________________________
   2.______________________________________________________________________________
   3.______________________________________________________________________________

6. Generate an item for use in an instrument to assess impact of instruction, including unintended outcomes:
7. Generate three items for use in an instrument to assess long-term effects:

Directions: ____________________________________________________________

_____________________________________________________________________

1. ________________________________________________________________

2. ________________________________________________________________

3. ________________________________________________________________
Exit Attitude

Opinion Survey*
(Example)

Directions: This opinionnaire attempts to assess attitudes of participants toward educational evaluation. Please indicate your agreement or disagreement with each statement by checking ( ) the answer that best describes how you personally feel, regardless of whether other people may agree or disagree with you. Although many of the questions may appear similar, please judge each one on an individual basis. Since we need to know your attitude, please answer each question frankly and honestly. There are no correct responses. There is no time limit, but respond as quickly as you can, and do not leave out any of the statements.

1. I see and feel deeply a need for increased evaluation information for decision-making in my field.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

2. Educational evaluation usually results in arbitrary judgements about the educative process.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

3. There is currently too much concern with evaluation in education.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

4. Educators do not receive sufficient training in evaluation.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

5. Educative processes are too complex to be evaluated objectively.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

6. Evaluation is an attempt to reduce education to a mechanistic process.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

7. Educational evaluation usually results in improvements in educational practice.

   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree
8. Evaluation should aid an educator in revising his goals even while the program is in progress.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

9. Evaluation interferes with the running of schools more than helps.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

10. Intuition and general knowledge of practitioners are more valuable than formal evaluative procedures in making decisions in education.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

11. Using educational evaluation in my work does not appeal to me.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

12. Money spent on evaluation contributes more to the improvement of education than any other expenditure.

<table>
<thead>
<tr>
<th>Strongly</th>
<th>Disagree</th>
<th>Undecided or Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

WORKSHOP OUTCOMES

Presentor: 4-E
Date: 

Workshop Title: Evaluation of Instruction

1. Please provide below a very brief description of the workshop as you experienced it, and as you would like to see it described to future participants.


2. Please list below what you perceive to be the main objectives of this workshop for participants. (Use back of page, if necessary.)

ATTAINMENT:  

OBJECTIVE

a b c  A. ________________________________________________________________

a b c  B. ________________________________________________________________

a b c  C. ________________________________________________________________

a b c  D. ________________________________________________________________

a b c  E. ________________________________________________________________

a b c  F. ________________________________________________________________

3. Please circle the capital letter (above) in front of each of the objectives that you believe were attained by you as a result of the workshop. Then indicate degree of attainment for each objective you circled. Indicate degree by circling the appropriate small letter.

(Scale: a = minimal, b = partial, c = great).
PRODUCT RATING

Please rate the Evaluation of Instruction Kit by circling your response to each statement. Support each choice with a brief statement of rationale. (Scale: SD = strongly disagree, D = disagree, U = undecided, A = agree, SA = strongly agree.)

1. This kit is easy for me to understand; the ideas have been well communicated. Rationale: SD D U A SA

2. This kit seems difficult to use. Rationale: SD D U A SA

3. I think that using this kit would significantly improve my instruction. Rationale: SD D U A SA

4. I will probably use at least parts of this kit. Rationale: SD D U A SA

Comments:
APPENDIX D:
1. End of Workshop
   b. Reports
Evaluation Report
(Teachers)

Purpose & Description:

A workshop on the evaluation of instruction was held at the Nisonger Center on June 11, 12, and 13. Participants represented staff from the Division of HR and PP, state institutions for NE, and the community college programs in Ohio. Participants were drawn from various program levels and interests. Instructors were drawn from the faculty of The Ohio State University. The workshop was designed to help participants develop skills for improving their classroom instruction. These skills consisted of developing and implementing an evaluation model.

Data Analysis:

Gestalt statements suggest that participants felt confused and frustrated the first day due to vagueness of instruction and jargon, but now feel they have some information to take home to utilize and integrate into their specific situation.

Instruments generated include 9 needs assessment and 8 skill, affect, knowledge evaluations. The format of the needs assessment include:
1. affective
2. cognitive
3. congruency
4. quality skill.

Participants felt that they partially attained the goal of generating an instrument to monitor the instructional process. The greatest degree of feeling of attainment or partial attainment was in generating an instrument for assessing skill, knowledge, affects, achievements or unintended outcomes.

Data Analysis of Topic Feedback Instruments:

1. Participants wanted to have expansion done in the areas of introduction, explanation of terminology and content, examples of practical application, prioritizing, and brainstorming skills.

2. The majority of participants did not feel anything should be deleted from the first day's session.

The most helpful format for the participants utilized lecture and discussion. Most agreed that the theoretical framework and specific suggestions and advice presented were useful and met individual needs. There was also agreement that the most important issues were addressed. Almost all participants felt the session's length was too long. A little over one half of the respondents noted an interest in participating in a short interview to work on the session's development.
Goals for Participants

Goals Attained

1. 3  G P P
2. 9  P P P P F F P G G
3. 5  P P P F

Instruments Generated

1. Needs Assessment  9 instruments
2. Skill, affect, knowledge evaluation  8 instruments
3. Assessing Impact
4. Analysis of Materials

Format of Needs Assessment Instrument

<table>
<thead>
<tr>
<th>Contingency</th>
<th>Affective</th>
<th>Cognitive</th>
<th>Congruency</th>
<th>Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Cognitive criteria

<table>
<thead>
<tr>
<th>time</th>
<th>frequency</th>
<th>quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

Gestalt Workshop

I appreciate

- individual help (1)
- adaptation to change, flexibility (12)
- ability to work on useable tool (2)
- very little (2)
- effort (1)
- handouts (2)
- communication (1)
- being able to select someone with common interests (1)
- informality (4)

I resent

- vague directions, confusion, jargon (7)
- the time it takes to develop evaluation (1)
- lack of formal instruction (1)
- lack of experience, failure to introduce early (1)
- heat (1)
- lack of structure (1)
- length of day (1)
- nothing (1)
- not knowing goals (1)
Gestalt Workshop
(continued)

**I want**
- help in evaluating curriculum for young children (1)
- more interaction with other participants (1)
- to get this over (1)
- sharing of tools (1)
- large group (1)
- to take a course in evaluation (3)
- clearer definition of terms (4)
- to apply workshop (5)

**I learned**
- I don't know enough basics (2)
- technique for group choosing of objectives (1)
- to develop instrument on affect (1)
- keeping abreast with evaluation is difficult (1)
- to expand tools (1)
- expanded concepts and knowledge concerning evaluation (9)
- very little (1)
- some information (2)

**I plan**
- to expand concepts (12)
- to struggle through (2)
- get sleep (1)
- define terms (1)
- be absent (1)

**I wish**
- We had more time (1)
- to know more information (11)
- discussion sessions (1)
- do more with specifics (1)
- resource handouts (2)
- days were shorter (1)
- I could write faster (1)
- not much (1)

**I feel**
- less frustration due to increased understanding (7)
- frustrated and confused workshop may not meet individual needs (4)
- sleepy (3)
- warm (1)
- uneasy about what I've accomplished (1)
- I've wasted time (1)
DEBRIEFING FORM

Instructor: "I would like to explore with you your responses on the instruments I have used in this workshop. After looking through the responses of the participants, I would like for you to help me solve the following problems."

1. Setting participants to understand terminology at the very beginning, as well as an explanation of the model. Student reaction and recommendations: a lack of knowledge of jargon led to participants' confusion about directions. The terminology should be defined at the very beginning, as well as the evaluation model being utilized.

2. Anxiety level of participants. Student reactions and recommendations: Participants felt a lack of clear-cut objectives due to their ignorance of terminology and evaluation model.

3. Effective Grouping -- participants' interests should have been decided before the workshop.

Adapted from Abdar. Viewpoints, July, 1972.
Evaluation Report (Program)

The results of our evaluation as of the second day of a three day workshop entitled "Evaluation of Instruction" indicate that partial fulfillment of goals for a majority of participants has been obtained. (see table 1).

<table>
<thead>
<tr>
<th>Goal</th>
<th>Great</th>
<th>Partial</th>
<th>Minimal</th>
<th>No-Stall</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Data was obtained from an evaluation instrument completed by participants entitled "Participant Goals" (see appendix).

A second evaluation instrument entitled "Gestalt Statements" appendix B, was administered to participants to elicit "gut level" reactions to the material presented and methods of presentation during this workshop. Statements were classified as positive, negative, and self-evaluatory. The following list is a summation of the content of the responses. (see table 2).

The third evaluation instrument entitled "topic feedback" (see appendix C) indicated that the participants generally agreed with the format, relevance, theoretical framework, suggestion and advice, major thrust of the first day of the workshop. The consensus of opinion was that the length of the first day was too long. (see table 3). Over one(1) of the questionnaire was concerned only with a single topic of the first day (the lectures of Jerry's) and is not reflected in the above table. (see appendix D for this interpreted data).
Recommendation -

1) Need for accurate pre-assessment of participants in terms of interest and what was expected for the workshop. Need for background information prior to attendance at workshop in terms of:
   a) terminology
   b) a basic reference article
   c) the dealing of the kit booklet "How to Use Evaluation of Instruction Kit"

2) Heavy emphasis should be placed on the explanation and use of the kit during the first morning session.

3) Need for model explanation early in the session so that later practice experience have greater meaning.

4) Practice experience should reflect the information given and occur after the presentation and review of pre-workshop material.

5) Flexibility, adaptability of group and resources both in terms of material and personal should be kept as in integral part of the workshop.

6) The one to one relationship between instructor and participant was seen as valuable and should be continued in future workshops.

7) The openness and honesty of instructor and participants increase learning.
8) Air-conditioning was too hot the first day and too cold the third day.

9) In general, a jolly good time was had by all and all felt positive about the experience.

Special thanks go to Jerry Adams for preparation, flexibility and general tolerance of a situation which called for these qualities.
Evaluation of Instruction Workshop
(Some comments from the developer)

The Evaluation of Instruction Workshop was designed to assist instructors of higher education and continuing education. The workshop was designed to facilitate use of the Evaluation of Instruction Kit. Twelve objectives of participants were stated for the workshop and a needs assessment of participants was done on the basis of the objectives.

An unusually large number of indicators are used to assess the impact of the kit and the workshop: participants are assessed on knowledge, achievements, skill, opinion, products, use, and comprehension; the workshop is assessed on goal achievement, topics, participant expectations, participant criteria, participant preferences and participant success; the kit is assessed on amount of use, user success, user opinion, and ratings by nationally recognized experts.

RESULTS

It is appropriate for the developer of a workshop to collect evaluation information for revision purposes; it is not appropriate for the developer alone to assess, interpret, and report the success of a workshop.

Several independent evaluation specialists will be employed to assess and report the success of the workshop. Some data are already available, however, and are included in the Appendix. No deletions or interpretations have been made in the data. Condensing and interpreting data on success are left for the evaluation specialists. Many more types of information were collected and will be available in the final report. The final report will include a follow-up study; completion is expected in mid-August.
CHARACTERIZATION OF WORKSHOP

Please provide below a very brief description of the workshop as you experienced it, and how you would like to see it described to future participants.

To develop your own evaluation material be it instruction, children or program based on the evaluation model supplied.

An overview of evaluation procedures and resources to assist one in implementing evaluation tools.

I started out in total frustration -- as our group recommended, I would have liked material in advance so that I could have studied it prior to coming. (It came at a poor time for me -- I'm still full of last quarter's trials).

A different experience. A good workshop to get a basic overview of evaluation.

A thought provoking workshop dealing with the entire process of evaluation.

The workshop is a good general introduction to evaluation. It is strongest in presentational information and material (with practical experience) in evaluation of instruction. I would like more information on program evaluation.

A workshop designed to provide the participants knowledge of an evaluation model, and opportunities to develop evaluation instruments for improving instruction.

A workshop to give you a basic approach to the development of a model of evaluation to fit your individual needs.

A workshop experience in which some of the fundamentals of evaluation were gained.
ENTRY ACHIEVEMENT *
(People Who Attended Workshop)

This is an average of the level of competency for the following categories that the people who attended the workshop came in with. (The rating scale is as follows: 1=no competency, 2=minimum competency, 3=moderate competency, 4=high competency, 5=very high competency)

<table>
<thead>
<tr>
<th>Category</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop instrument to logically analyze instructional materials</td>
<td>2.1</td>
</tr>
<tr>
<td>Develop needs assessment instrument</td>
<td>2.2</td>
</tr>
<tr>
<td>Develop criterion referenced test</td>
<td>2.0</td>
</tr>
<tr>
<td>Develop instrument to monitor the instructional process</td>
<td>2.1</td>
</tr>
<tr>
<td>Develop an instrument to assess impact, including unintended outcomes</td>
<td>1.9</td>
</tr>
<tr>
<td>Develop instrument to collect reactions from participants</td>
<td>2.3</td>
</tr>
<tr>
<td>Develop an instrument to assess long-term effects</td>
<td>1.8</td>
</tr>
<tr>
<td>Develop evaluation design</td>
<td>1.9</td>
</tr>
<tr>
<td>Find needed resources for data analysis</td>
<td>1.8</td>
</tr>
<tr>
<td>Validate an instrument</td>
<td>1.6</td>
</tr>
<tr>
<td>Interpret evaluation results</td>
<td>2.3</td>
</tr>
<tr>
<td>Report evaluation results</td>
<td>2.3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2.0</td>
</tr>
</tbody>
</table>

1=No Competency. Respondent has had no training and experience base from which to work
2=Minimum Competency. Respondent has had some training or experience in this area.
3=Moderate Competency. Respondent has fairly thorough training or experience in this area and could perform the task moderately well.
4=High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.
5=Very High Competency. Respondent is completely confident in performing this task.

*Self Report
**EXIT ACHIEVEMENT**
*(People Who Attended Workshop)*

This is an average of the level of competency for the following categories that the people who attended the workshop left with. (The rating scale is as follows: 1 = no competency, 2 = minimum competency, 3 = moderate competency, 4 = high competency, 5 = very high competency)

<table>
<thead>
<tr>
<th>Category</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop instrument to logically analyze instructional materials</td>
<td>2.6</td>
</tr>
<tr>
<td>Develop needs assessment instrument</td>
<td>2.8</td>
</tr>
<tr>
<td>Develop criterion referenced test</td>
<td>2.1</td>
</tr>
<tr>
<td>Develop instrument to monitor the instructional process</td>
<td>2.7</td>
</tr>
<tr>
<td>Develop instrument to collect student reactions</td>
<td>2.7</td>
</tr>
<tr>
<td>Develop instrument to assess the impact of instruction, including unintended outcomes</td>
<td>2.4</td>
</tr>
<tr>
<td>Develop instrument to assess the long-term effects of instruction</td>
<td>2.1</td>
</tr>
<tr>
<td>Develop evaluation design for instruction</td>
<td>2.7**</td>
</tr>
<tr>
<td>Find needed resources for data analysis</td>
<td>2.8</td>
</tr>
<tr>
<td>Validate an instrument</td>
<td>1.9</td>
</tr>
<tr>
<td>Interpret evaluation results</td>
<td>2.2</td>
</tr>
<tr>
<td>Report evaluation results</td>
<td>2.8</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>2.5</td>
</tr>
</tbody>
</table>

*Self Report

** A major participant interest
The rating scale used for this is as follows: SD=strongly disagree, D=disagree, U=undecided, A=agree, SA=strongly agree.

1. This kit is easy for me to understand; the ideas have been well communicated.
   FREQUENCY of responses: SD=0, D=1, U=2, A=6, SA=0
   Agree - The format of organization is concise and provides good organizational management skills.
   Agree - Once I understood terminology.
   Undecided - I still need to work with it alot more so that I have a better idea of what I'm doing.
   Agree - The Ideas were not always well communicated.
   Comments
   Undecided - Need further study of kit
   Agree - I understand now the outline and how I could use it but short form and with additional questions if necessary.
   Agree - Only after I understood terminology and evaluation model.

2. This kit seems difficult to use.
   FREQUENCY of responses: SD=2, D=6, U=1, A=0, SA=0
   Disagree - Once I really get into it I don't think I'll have problems with it.
   Disagree - Once I understood an evaluation model.
   Strongly Disagree - The format of organization is concise and provides good organizational management skills.
   Strongly Disagree -
   Comments
   Undecided - Need further study of kit.
   Disagree - Spoon-feeding.
   Disagree - It's well organized and has a manual that explains it.
   Disagree - Not at all

3. I think that using this kit would significantly improve my instruction.
   FREQUENCY of responses: SD=0, D=1, U=0, A=5, SA=3
   Strongly Agree - I will have ready references to fall back on so that I won't need to rely on my inadequate memory.
   Agree - If it is an evaluation model is used program-wide
   Strongly Agree - I plan to study the booklet and kit thoroughly and put it to use.
   Agree - I don't instruct.
   Comments
   Agree - The basic format appears to be very usable for my situation.
   Strongly Agree - I have just used a blank sheet of paper in the past and this could include specifics I want to have.
   Agree - My instruction would be better organized easier to evaluate and have more accountability.
   Agree - I hope the time element does not get in the way and hamper the kit's use.
4. I will probably use at least parts of this kit (continued).

Strongly Agree - Next week will find me actively delving into the material.
Agree - It is something our program has never had.

Strongly Agree - I would like to get the impact it could give me to improve my teaching.
Strongly Agree - It will be helpful in evaluating my work.
Agree - This will probably be my first resource when I start developing a model and/or an evaluation instrument.

General Comments:

I work best when not interrupted by group activities. Much of this workshop was used to become acquainted with materials and tools and the study-learning will take place next week. I hope to be able to contact you and become efficient in this technique.

Just a great job.

I'm pleased with what I've learned. I intend to use the kit to add more validity to my instruction and program.

I really appreciate being given such a tool. So many times you attend a workshop and go home with notes. When you go back over them sometimes you can't recall things you need to recall to try something new. This kit bridges this gap.
APPENDIX E:
1. After the Workshop
THE OHIO STATE UNIVERSITY

June 18, 1973

Jerry Adams
439 Nisonger Center
McCandless Hall
1580 Cannon Drive
Columbus, Ohio 43210

Dear,

I hope your evaluation activities have been satisfactory for you since attending the evaluation workshop.

An evaluation specialist will call you (9 A.M. to 4 P.M.) between July 2 and July 13. The specialist will ask you questions about your use of evaluation and success in evaluation as a result of the workshop.

The interview would be much easier if you would assemble a copy of all that you developed up to July 2, so that when the call comes you can talk about your products easily. Also, assembling a copy of each of your creations will make sending us a copy (for rating and critiquing) after the call much easier. Please send us the copies immediately after the call.

Thank you.

Sincerely,

Jerry Adams

JA/rbk
enc.: self rating achievement
Please fill this out before July 2nd.

FOLLOW-UP ACHIEVEMENT (SELF-REPORT)

Please assist the planning of the next Evaluation of Instruction Workshop.

Circle the number that best describes your level of competency for each of the following items. The responses are divided into five categories:
1=No Competency. Respondent has had no training and experience base from which to work.
2=Minimum Competency. Respondent has had some training or experience in this area.
3=Moderate Competency. Respondent has fairly thorough training or experience in this area and could perform the task moderately well.
4=High Competency. Respondent has a good background in this area and is completely comfortable in performing this task.
5=Very High Competency. Respondent is completely confident in performing this task and has done creative work in this area.

1. Develop instrument to logically analyze instructional materials............................................ 1 2 3 4 5
2. Develop needs assessment instrument.............................................1 2 3 4 5
3. Develop criterion referenced test.............................................1 2 3 4 5
4. Develop instrument to monitor the instructional process.......1 2 3 4 5
5. Develop instrument to collect student reactions..............1 2 3 4 5
6. Develop instrument to assess the impact of instruction, including unintended outcomes..........................1 2 3 4 5
7. Develop instrument to assess the long-term effects of instruction............................................. 1 2 3 4 5
8. Develop evaluation design for instruction..........................1 2 3 4 5
9. Find needed resources for data analysis..........................1 2 3 4 5
10. Validate an instrument............................................................1 2 3 4 5
11. Interpret evaluation results.................................................... 1 2 3 4 5
12. Report evaluation results.......................................................... 1 2 3 4 5
8. Have you requested the services of an evaluation specialist as a result of the workshop? If no, go to #10.

9. Do you feel that you were more efficient in asking for service as a result of ideas you gained from the workshop or the kit?

10. Has attending the workshop resulted in any major benefits for you that were not part of the workshop objectives? If yes, please elaborate:

11. Did attending the workshop result in major problems for you? If yes, please elaborate:

12. Please read out your scores on Exit Achievement (self report):

1____ 2____ 3____ 4____ 5____ 6____ 7____ 8____ 9____ 10____

11____ 12____.

Thank you for your time and information. If you will send in the instruments and design you now have, we will rate them and critique them. These should be mailed to: Jerry Adams
Room # 439 Nisonger Center
McC Campbell Hall
1580 Cannon Drive
Columbus, Ohio 43210

If you want additional consulting concerning evaluation of instruction, please feel free to call Jerry at 614-422-9780. He will be glad to provide assistance.
Date ______________________

TELEPHONE WORKSHOP INTERVIEW

Introduction:

Hello, I am Helen Armstrong. I am calling regarding the Evaluation of Instruction Workshop you recently attended at Misonger Center.

Name ________________________ Code __________

1. Please list all instruments you have used or constructed as a direct result of the workshop and/or kit (including during and after the workshop)?

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Rating</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Please list any instruments you used or constructed that are based on the kit alone and not the workshop?

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Rating</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Please rate each of the above instruments as "good" "medium" or "poor" (read each out and get a rating).

4. How many of the instruments (in the kit or developed from the workshop or kit) have you already used? ______
Which ones (if any)? (Place a check by each.)

5. How many do you have specific plans to use in the near (few months) future? ______ Which ones (if any)?
(Place a double check by each.)

YES      NO

____  ____  6. Do you now have an overall evaluation
design? If yes, please outline components.

____  ____  Did you use the kit in the development of
your design?

____  ____  If you do not already have an evaluation
design, are you using the kit to work on an
overall evaluation design?

____  ____  7. Since the workshop have you developed
or used instruments or an evaluation design
independent of the workshop and the kit? If
so, please identify:
____  ____  If not, would you have in the time since you attended the workshop if you had not attended the workshop and had not received the kit?

____  ____  8. Have you requested the services of an evaluation specialist as a result of the workshop? If no, go to #10.

____  ____  9. Do you feel that you were more efficient in asking for service as a result of ideas you gained from the workshop or the kit?

____  ____  10. Has attending the workshop resulted in any major benefits for you that were not part of the workshop objectives? If yes, please elaborate:

____  ____  11. Did attending the workshop result in major problems for you? If yes, please elaborate:

12. Please read out your scores on Exit Achievement (self-report):

1    2    3    4    5    6
7    8    9    10   11   12
Thank you for your time and information. If you will send in the instruments and design you now have, we will rate them and critique them. These should be mailed to:

Jerry Adams  
439 Nisonger Center  
McCampbell Hall  
1580 Cannon Drive  
Columbus, Ohio 43210

If you want additional consulting concerning evaluation of instruction, please feel free to call Jerry at 614-422-9780. He will be glad to provide assistance.
APPENDIX E:
2. After Receiving the Mailed Kit
THE OHIO STATE UNIVERSITY

June 18, 1973

Jerry Adams
439 Nisonger Center
McCormick Hall
1580 Cannon Drive
Columbus, Ohio 43210

Dear

I hope your evaluation activities have been satisfactory for you since you received your Evaluation of Instruction kit.

An evaluation specialist will call you (9 A.M. to 4 P.M.) between July 2 and July 13. The specialist will ask you questions about your use of evaluation and success in evaluation as a result of the workshop.

The interview would be much easier if you would assemble a copy of all that you developed up to July 2, so that when the call comes you can talk about your products easily. Also, assembling a copy of each of your creations will make sending us a copy (for rating and critiquing) after the call much easier. Please send us the copies immediately after the call.

Thank you.

Sincerely,

Jerry Adams

JA/rbk
enc.: self rating achievement
Code no. __________

EXIT ATTITUDE

Opinion Survey*
(Example)

Directions: This opinionnaire attempts to assess attitudes of participants toward educational evaluation. Please indicate your agreement or disagreement with each statement by checking ( ) the answer that best describes how you personally feel, regardless of whether other people may agree or disagree with you. Although many of the questions may appear similar, please judge each one on an individual basis. Since we need to know your attitude, please answer each question frankly and honestly. There are no correct responses. There is no time limit, but respond as quickly as you can. And do not leave out any of the statements.

1. I see and feel deeply a need for increased evaluation information for decision-making in my field.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

2. Educational evaluation usually results in arbitrary judgments about the educative process.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

3. There is currently too much concern with evaluation in education.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

4. Educators do not receive sufficient training in evaluation.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

5. Educative processes are too complex to be evaluated objectively.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

6. Evaluation is an attempt to reduce education to a mechanistic process.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree

7. Educational evaluation usually results in improvements in educational practice.
   ( ) Strongly Disagree ( ) Undecided or Neutral ( ) Agree ( ) Strongly Agree
8. Evaluation should aid an educator in revising his goals even while the program is in progress.

( ) Strongly Disagree ( ) Undecided or Agree
Disagree Neutral Strongly Agree

9. Evaluation interferes with the running of schools more than helps.

( ) Strongly Disagree ( ) Undecided or Agree
Disagree Neutral Strongly Agree

10. Intuition and general knowledge of practitioners are more valuable than formal evaluative procedures in making decisions in education.

( ) Strongly Disagree ( ) Undecided or Agree
Disagree Neutral Strongly Agree

11. Using educational evaluation in my work does not appeal to me.

( ) Strongly Disagree ( ) Undecided or Agree
Disagree Neutral Strongly Agree

12. Money spent on evaluation contributes more to the improvement of education than any other expenditure.

( ) Strongly Disagree ( ) Undecided or Agree
Disagree Neutral Strongly Agree

PRODUCT RATINGS

Please rate the Evaluation of Instruction Kit by circling your response to each statement. Support each choice with a brief statement of rationale. (Scale: SD = Strongly disagree, D = disagree, U = undecided, A = agree, SA = strongly agree.)

1. This kit is easy for me to understand; the ideas have been well communicated. SD  D  U  A  SA
   Rationale:

2. This kit seems difficult to use. SD  D  U  A  SA
   Rationale:

3. I think that using this kit would significantly improve my instruction. SD  D  U  A  SA
   Rationale:

4. I will probably use at least parts of this kit. SD  D  U  A  SA
   Rationale:

Comments:
TELEPHONE INTERVIEW FOR KIT RECIPIENTS

Introduction: Hello, I am Helen Armstrong. I am calling regarding the Evaluation of Instruction Kit you received in the mail.

Name ____________________ Code __________

YES  NO

1. Did you receive the kit? (If no, "thank you and good-bye.") On what date (approximately) did you receive the kit?


2. Was the kit damaged?

If yes, did the damage affect your use of the kit? Please explain:


3. Have you examined the kit?

If no, do you intend to examine the kit?
If so, when?

May I call you back after that date?


4. Have you used or constructed any instruments as a result of the kit?
If so, please list them:

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Rating</th>
<th>Use</th>
</tr>
</thead>
</table>

5. Please rate each of the above instruments as "good," "medium," or "poor." (Read each out and get a rating.)

6. How many of the instruments (in the kit or developed from the kit) have you already used? __________________
Which ones (if any)? (Place a check by each.)

7. How many (from #4) do you have specific plans to use in the near (few months) future? __________________
Which ones (if any)? (Place a double check by each.)

YES  NO

8. Do you now have an overall evaluation desing? If yes, please outline components:

Did you use the kit in the development of your design?
YES  NO

If you do not already have an evaluation design, are you using the kit work on an overall evaluation design?

___  ___  9. Since receiving the kit have you developed or used instruments or an evaluation design *independent* of the kit? If so, please identify:

___  ___  If not, would you have in the time since you received the kit *if* you had not received the kit?

___  ___  10. Have you requested the services of an evaluation specialist as a result of examining the kit? If no, go to #12.

___  ___  11. Do you feel that you were more efficient in asking for services as a result of the ideas you gained from the kit?

___  ___  12. Has receiving the kit resulted in any unplanned benefits for you? If yes, please elaborate:
13. Please read out your scores on Exit Achievement (Self-Report):

1 ____  7 ____ 
2 ____  8 ____ 
3 ____  9 ____ 
4 ____ 10 ____ 
5 ____ 11 ____ 
6 ____ 12 ____ 

Please send us your ratings on the kit and on the exit opinionnaire (if appropriate) as well as the instruments and/or design you now have. We will rate them and critique them right away. These should be mailed to:

Jerry Adams
439 Nisonger Center
1580 Cannon Drive
The Ohio State University
Columbus, Ohio 43210

If you would like some free consulting service concerning evaluation of instruction, please feel free to call Jerry Adams at 614-422-9780. He will be glad to provide assistance.

Thank you for your time and information.
APPENDIX F:
1. Instrument for Assessment of Workshop Instruments
2. Instrument for Assessment of Kit Instruments
3. Instrument for Assessment of Kit (by Experts)
EVALUATION PRODUCTS

Please rate the instruments below by circling the appropriate number in front of each item. Use the following criteria:

1 - **Poor** (so inadequate that it should not be used; and/or needs much improvement to be worthwhile but is perhaps salvageable.)

2 - **Mediocre** (not outstanding; uncertain of value; and/or usable)

3 - **Good** (some outstanding merit; well conceived; efficient; clear)

**EVALUATION INSTRUMENTS**

(Used in Workshop)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Background Information</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Achievement Criteria</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Criteria for Participants</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Entry Skill</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Entry Attitude</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Entry Knowledge</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Group Effectiveness Scale</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Exit Knowledge</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Product Rating</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Exit Achievement</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Exit Skill</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>Workshop Outcomes</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Topic Feedback</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Goals for Participants</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Debriefing Form</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Objectives for Session</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Content Validation</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Telephone Workshop Interview</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Telephone Interview for Kit Recipients</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Instrument Rating Scale</td>
</tr>
</tbody>
</table>
EVALUATION PRODUCTS

Please rate the instruments below by circling the appropriate number in front of each item. Use the following criteria:

1 - Poor (so inadequate that it should not be used; and/or needs much improvement to be worthwhile but is perhaps salvageable.)

2 - Mediocre (not outstanding; uncertain of value; and/or usable)

3 - Good (some outstanding merit; well conceived; efficient; clear)

EVALUATION INSTRUMENTS

1 2 3 Analysis of Materials
1 2 3 Product Rating
1 2 3 Observation Scale
1 2 3 Entry Achievement
1 2 3 Student Criteria
1 2 3 Entry Skill
1 2 3 Entry Knowledge
1 2 3 Entry Attitude
1 2 3 Background Information
1 2 3 Gestalt Statements
1 2 3 Group Effectiveness Scale
1 2 3 Group Effectiveness Observation Form
1 2 3 Comprehensive of Student Involvement
1 2 3 Goals for Session
1 2 3 Topic Feedback
1 2 3 Open Ended Responses to Session
1 2 3 Debriefing Form
1 2 3 Workshop Outcomes
1 2 3 Exit Attitude
1 2 3 Exit Achievement
1 2 3 Uses of Format
Interview Guide
Product Rating
Follow-up Achievement
Follow-up Attitude
Follow-up Rating - Use
Follow-up Rating - Organizational Context
THE OHIO STATE UNIVERSITY

June 26, 1973

Jerry Adams
Rm 439 Nisonger Center
1580 Cannon Drive
The Ohio State University
Columbus, Ohio 43210

Basic letter to experts; each letter individualized

Dear

Since you are recognized as a leader in educational evaluation, I would appreciate your assistance in assessing the content validity of the enclosed kit. The kit was designed for use by instructors of higher education and continuing education. It is hoped that the kit will facilitate the development of an overview of instructional evaluation; it is also hoped that instructors will use or adapt some of the instruments for evaluation of their instruction. The instructors will receive the kit (now in manual form) in a letter file box with folders identifying the different categories. Storage space will be available in the box for a 247-item bibliography, rough drafts, their own design and instruments, and data collected from the instruments.

Please read the introductory booklet, examine the manual, and fill out the rating form. The rating needs to be returned on or before July 16th. (Less than 30 minutes should be required to go through the entire process.)
Thank you in advance for your cooperation.

Sincerely,

Jerry Adams

JA/rac

Enclosures:  A form for Rating Evaluation of Instruction
             An introductory booklet on "How to Use
             Evaluation of Instruction Kit"
             A kit (in manual form) for Evaluation of
             Instruction
Rating Evaluation of Instruction

Please rate Evaluation of Instruction by circling the appropriate number in front of each statement using the following criteria:

1. Very Poor = (so inadequate that it should be discontinued in terms of this criterion.)
2. Poor = (needs much improvement to be worthwhile but is perhaps salvageable.)
3. Average = (not outstanding, mediocre, uncertain of value.)
4. Good = (some outstanding merit on this criterion.)
5. Very Good = (definitely outstanding on this criterion.)

1 2 3 4 5 Importance of the concept of the kit relative to other developments you are aware of in evaluation of instruction. Rationale:

1 2 3 4 5 Importance of the information selected for the kit relative to needs you perceive for the intended audience. Rationale:

1 2 3 4 5 Ease to use you would project or estimate for intended audience. Rationale:

1 2 3 4 5 Helpfulness you would project or estimate for intended audiences. Rationale:

1 2 3 4 5 Overall quality of the kit. Rationale:

Faulty information in the kit:

General Comments: