

1-1-1975

Skill identification and subsequent development of training materials for the Fortune-Hutchinson Evaluation Methodology.

Virginia Percy Mitchell
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_1

Recommended Citation

Mitchell, Virginia Percy, "Skill identification and subsequent development of training materials for the Fortune-Hutchinson Evaluation Methodology." (1975). *Doctoral Dissertations 1896 - February 2014*. 2985.
https://scholarworks.umass.edu/dissertations_1/2985

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

UMASS/AMHERST



312066013544765

SKILL IDENTIFICATION AND SUBSEQUENT DEVELOPMENT OF
TRAINING MATERIALS FOR THE FORTUNE-HUTCHINSON
EVALUATION METHODOLOGY

A Dissertation Presented

By

VIRGINIA PERCY MITCHELL

Submitted to the Graduate School of the
University of Massachusetts
in partial fulfillment of the requirements
for the degree of

DOCTOR OF EDUCATION

June

1975

Educational Evaluation

(c) Virginia Percy Mitchell
All Rights Reserved

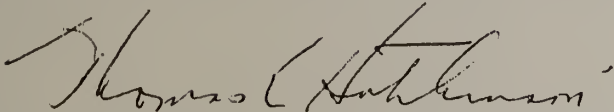
SKILL IDENTIFICATION AND SUBSEQUENT DEVELOPMENT OF
TRAINING MATERIALS FOR THE FORTUNE-HUTCHINSON
EVALUATION METHODOLOGY

A Dissertation

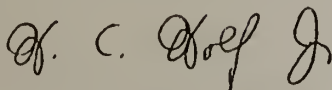
By

VIRGINIA PERCY MITCHELL

Approved as to style and content by:



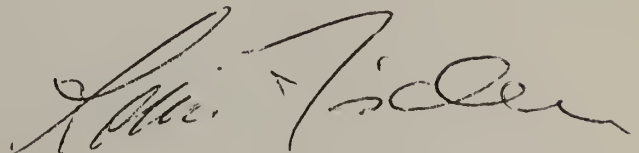
Dr. Thomas E. Hutchinson, Chairman



Dr. William C. Wolf, Jr., Member



Dr. Maurianne Adams, Member



Dr. Louis Fischer, Acting Dean
School of Education

June 1975

ACKNOWLEDGMENTS

The development and completion of this work would not have been possible without the support and encouragement of many individuals. Paula Nassif, Larry Benedict, and Philip Brooks have been invaluable from the beginning when the topic for investigation had not even been specified. They helped me to narrow down some alternatives and continued to give generously of their time throughout the course of the work which eventually took shape. My thanks also to Sister Ann Luciano and William Welch who helped me define the more complex skill of the two on which I concentrated.

Since a third of this work concerned the field testing of instructional materials, I'd like to thank my field test participants who, though busy, took the time to contribute to this effort. Many thanks to Steve Guild, Larry Marcus, Charles Adams, Robert Welch, Michael Beaudoin, Kathryn Girard, Elizabeth Kipp, Robert Powers, Gloria Figueroa-de-Guevara, Parnell Hegerman, Sylvia Cuomo, David Bristol, and the members of Dr. Hutchinson's "Evaluation Practicum."

I'd like to thank the members of my committee, Dr. William Wolf, Dr. Mary Quilling, and Dr. Maurianne Adams for their acute comments and valuable suggestions for the improvement of this work. I'd also like to thank Dr. William Gorth for his advice and information concerning available training materials for evaluators.

My special thanks go to my advisor, Dr. Thomas Hutchinson, who helped me from the beginning to formulate the problem to be investigated and to focus my thinking. He virtually furnished the subject matter which formed the basis for this work for, without him, the Fortune-Hutchinson Evaluation Methodology would not exist. Dr. Hutchinson's unique ability to specify the steps necessary to attack a problem in the most appropriate and expeditious way has been invaluable to me.

Finally, my thanks go to my husband Frank who sustained me, even in my worst moods, and kept me from letting this work become all-consuming.

To Frank
and
To My Mother

ABSTRACT

Skill Identification and Subsequent Development of
Training Materials for the Fortune-Hutchinson
Evaluation Methodology (June 1975)
Virginia Percy Mitchell, B.A., M.Ed.
University of Massachusetts

Directed by: Dr. Thomas E. Hutchinson

That the field of education is in need of usable evaluation tools has been frequently documented (Scriven, 1967; Glass, 1969). Further documentation has been presented to describe the potential of the Fortune-Hutchinson Evaluation Methodology (F-H) for meeting this need (Gordon, 1973; Benedict, 1973). The problem presented in this dissertation was that of improving the means by which a future evaluator learns how to use F-H.

At the time of this writing, F-H evaluators were trained by means of the most current version of the documented methodology which was presented, discussed, and eventually applied in an evaluation setting. Because of indications that improvements could be made in these methods (Gordon, 1973; Benedict, 1973; Rosen, 1974), this author identified the need for a list of the skills an F-H evaluator needs to have in order of importance, and materials for training future evaluators in these skills.

The development of the skills list involved a search of the following sources: the most recent documentation of F-H; brief questionnaires distributed to graduate students who had had experience using F-H; dissertations concerned with F-H; and other sources such as other evaluation models and related literature. The final list was put in a priority order by this author, Dr. Hutchinson, and Dr. Benedict, using as the criterion for each skill, "importance for being a successful F-H evaluator."

The top two skills were chosen for the development of training materials. These were "The evaluator should be able to define evaluation," and "The evaluator should be able to deal with a lack of decision maker cooperation." Learning materials were developed according to various guidelines, primarily Tyler (1950), involving the development of instructional objectives and content material and criterion measures for these objectives. The mode chosen for both skills was the self-instructional workbook, as the one most consistent with available resources and the author's present abilities and knowledge. Both workbooks were field tested in order to identify any major problems with them or with the criterion tests given to determine the subjects' achievement of the instructional objectives.

The field test group for the workbook "Defining evaluation" consisted of individuals who had demonstrated interest in educational evaluation, and, for "Dealing with a lack of

decision maker cooperation," of those who had taken a course in evaluation methodology. These groups also took the criterion measure of competency in the skills, and responded to questions about their background and reactions. In addition, a group of four "experts" who had had all available classroom training in F-H, as well as having used F-H in evaluations, took the criterion measures, without the instruction, for both skills to document the necessity of the training packages.

The responses of field test participants to the review questions in both workbooks, their responses to the criterion measures, and their reactions to the materials were analyzed in terms of implications for the improvement of both workbooks. The field tests were successful in the sense that problems were identified in the workbooks, review questions, and the criterion measures, allowing for future revisions and subsequent field testing. As expected, the "expert" group could have profited from both workbooks, and especially from the second.

Specific results for the first workbook suggested a re-evaluation of competency criteria for one question, a modification of two instructional objectives, and some additions to the workbook itself. Results for the second indicated that some of the problems given in the workbook which an evaluator might have with insufficient decision maker cooperation should be reexamined for validity, mutual exclusivity, and clarity.

TABLE OF CONTENTS

Chapter	Page
I SKILL IDENTIFICATION AND THE TRAINING OF EVALUATORS.....	1
The Need for Training Evaluators.....	1
Some Existing Evaluation Training Materials.....	3
The Methodological Approach.....	10
The Fortune-Hutchinson Evaluation Methodology.....	11
The nature of F-H.....	12
Major steps of F-H.....	13
The Importance of the Fortune-Hutchinson Evaluation Methodology.....	22
The Training of Fortune-Hutchinson Evaluators.....	25
Statement of the Problem.....	27
General Description of Procedures.....	27
Some General Principles for Developing and Evaluating Training Materials.....	28
Evaluating effectiveness of learning experiences.....	35
Skill Identification: The First Step.....	38
II GENERAL DESIGN OF THE STUDY AND THE SPECIFIC DESIGN OF THE SKILL IDENTIFICATION PROCESS.....	41
General Design of the Study.....	41
Specific Design of Skill Identification Procedures.....	42
Search of F-H documentation.....	43
User questionnaires.....	44
Search of related dissertations.....	44
Search of other evaluation models.....	45
Putting skills list in priority order.....	46
Specifying skills on the list.....	48
III RESULTS OF THE SKILL IDENTIFICATION PROCEDURES.....	49
List of Skills for F-H in Final Priority Order....	50
Discussion of the Skills List.....	61
Methodological Inadequacies.....	65
Implications for Further Research.....	64
IV DESIGN OF TREATMENTS AND TESTS FOR SELECTED SKILLS..	68
Skill Selection.....	68
General Approach to the Design of Treatments and Tests.....	69
Selected Mode of Instruction.....	69

Chapter	Page
Design of Treatment and Test for "Defining Evaluation".....	71
Objectives.....	71
General structure of the booklet.....	72
General content of the booklet.....	73
Review questions.....	74
Final review.....	75
Facilitator's handbook.....	76
Design of Treatment and Test for "Dealing with a Lack of Decision Maker Cooperation".....	77
Defining the skill.....	77
Objectives.....	79
General structure of the booklet.....	80
General content of the booklet.....	81
Review questions.....	82
Final review.....	82
Facilitator's handbook.....	84
V FIELD TESTS OF MATERIALS FOR SELECTED SKILLS.....	85
Design of the Field Test for "Defining Evaluation".....	85
Background questions.....	87
Student reaction.....	88
Experts.....	89
Results and Discussion of the Field Test of "Defining Evaluation".....	89
Review questions.....	90
Final review.....	92
Motivation.....	104
The role of background, training, and experience.....	106
Student reaction.....	110
Other recommendations.....	114
The necessity of the workbook.....	115
Final comparisons.....	116
Design of the Field Test for "Dealing with a Lack of Decision Maker Cooperation".....	117
Background questions and student reaction....	119
Experts.....	119
Results and Discussion of the Field Test of "Dealing with a Lack of Decision Maker Cooperation".....	120
Review questions.....	120
Final review.....	122
Motivation.....	129
The role of background, training, and experience.....	133

Chapter	Page
Student reaction.....	134
Other recommendations.....	137
The necessity of the workbook.....	140
Final comparisons.....	141
 VI SUMMARY, CONCLUSIONS, AND IMPLICATIONS FOR FURTHER RESEARCH.....	 143
Summary.....	143
Conclusions.....	146
Implications for Further Research.....	148
 REFERENCES.....	 153
 APPENDICES.....	 157
 Appendix A: Evaluation Handbook #1, <u>Defining</u> <u>Evaluation</u> ; Final Review; Facilitator's Handbook to <u>Defining Evaluation</u>	 158
 Appendix B: Evaluation Handbook #2, <u>Dealing with a</u> <u>Lack of Decision Maker Cooperation</u> ; Final Review; Facilitator's Handbook to <u>Dealing with a Lack of Decision</u> <u>Maker Cooperation</u>	 159
 Appendix C: Background Questions; Student Reaction to the Workbook, "Defining Evaluation"; Student Reaction to the Workbook "Dealing with a Lack of Decision Maker Cooperation".....	 160

LIST OF TABLES

Table	Page
I Number and Percentage of Skills on Final List from Each Source by Category and Total.....	63
II "Defining Evaluation": Final Review Question Competency Achieved by the Field Test Group and the "Experts".....	93
III "Defining Evaluation": Scores on the Final Review and Related Variables.....	105
IV Analysis of Variance, Group Analysis.....	107
V T-Tests for Field Test of "Defining Evaluation"....	108
VI "Dealing with a Lack of Decision Maker Coopera- tion": Final Review Question Competency Achieved by the Field Test Group and the "Experts".....	123
VII "Dealing with a Lack of Decision Maker Coopera- tion": Scores on the Final Review and Related Variables.....	131
VIII T-Test for Field Test of "Dealing with a Lack of Decision Maker Cooperation".....	132
IX Results and Implications for Revisions for both Sets of Materials.....	138

C H A P T E R I

SKILL IDENTIFICATION AND THE TRAINING OF EVALUATORS

The Need for Training Evaluators

The state of evaluation methodology is sadly lacking, as has been well documented by Gordon (1973), Benedict (1973b) and others. If this is true, though, the state of available training materials for teaching evaluators the existing methodology is even more sparse. Stake and Denny (1969) put it this way:

In spite of the fact that evaluation is a desirable and often mandatory responsibility within funded programs, few schools of education provide explicitly relevant course work and supervised experience (p. 372).

They go on to say that the training programs from the laboratories across the country ". . . have yet to draw effectively upon the experience and training materials of each other" (p. 374). Thus, not only do the laboratories lack a common evaluation methodology, making the development of uniform training materials more difficult, but the diffusion of existing materials is poor.

With respect to the content of the training materials, Stake and Denny (1969) suggest a number of skills which would be useful for an evaluator to have. These include skills in "devising techniques and constructing instruments idiosyncratic to the evaluation tasks at hand . . ." as well as skills in ". . . training other professionals and paraprofessionals

in the use of such instruments and techniques . . ." (p. 374). The evaluator should understand the use of unobtrusive measures, as well as more traditional measures, and the fallibility of tests in general. Finally, Stake and Denny (1969) suggest that, in the selection and training of evaluators, "consideration should be given to their tolerance for ambiguity and to their ability to persevere in working on unpleasant tasks" (p. 375). While not a very optimistic view of educational evaluation, this suggests that the evaluator needs to have developed certain personality traits for success in this field.

Guba and Gephart (1970) agree that there is a need for more training materials and offer reasons for the existing lack.

While there are demands for educational development specialists, program implementors, dissemination specialists, evaluation experts, etc., existing training programs are not responsive to the need to train personnel for such roles. This is largely because the roles of such personnel are quite obscure--very little is known about them. Insofar as practitioners of such roles now exist, hard data are not available to describe what individuals in these roles do or how their efforts relate to the efforts of other roles, either newly emerging or old" (pp. 3-4).

They add, "There is clearly a shortage of effective training materials and of qualified trainers for use in RDD&E training programs" (p. 4).

Worthen and Sanders (1973) feel that in training evaluators, as opposed to training researchers, content from several disciplines should be sampled so that the evaluator can

become sensitive to a wide range of phenomena. Such disciplines might include psychology, sociology, and the other social sciences. The emphasis in training, they feel, should be on statistical analysis, measurement and psychometrics, survey research methods, and experimental design. They add that evaluators should also be trained in the techniques of inquiry and should have a breadth of practical experience in a variety of settings so that they might learn the constraints of contemporary schools. In this author's opinion, these recommendations would lead to the training of frustrated researchers, rather than evaluators, i.e., individuals who would understand the goals and methods of research and the limitations imposed by the context of existing programs to be evaluated. No special perspectives for the evaluator are suggested by Worthen and Sanders (1973). This type of training would lead to the inappropriate application of research tools to an evaluation situation which would be to the detriment of the program or enterprise and the frustration of the "evaluator."

Some Existing Evaluation Training Materials

If the word "evaluation" is interpreted broadly, there have been several efforts in the area of developing evaluation training tools which can be cited here. In the area of instructional objectives, Mager (1962) and Popham and Baker (1970) have written short training books on the subject.

Mager's book, Preparing Instructional Objectives, is self-instructional in nature and is designed for "anyone interested in transmitting skills and knowledge to others" (Mager, 1962). Establishing Instructional Goals, by Popham and Baker (1970), is supposed to help teachers decide what objectives they should attempt to achieve. These objectives would be judged by the adequacy of student accomplishment. This achievement may be cognitive, attitudinal, or psychomotor. The book also aimed at helping teachers decide which instructional activities to include in a teaching sequence and whether or not the sequence was effective.

For the instructor who has never thought in terms of specifying the behavior changes that (s)he would like to see, these books can be extremely valuable. (Note: Accepted English grammar permits the use of the pronoun "he" to pertain to either a male or female individual. This has often been reversed, however, and the use of "she" been adopted in reference to certain classes of employees, e.g. nurses, teachers, secretaries. This author therefore prefers to use "(s)he" in order to avoid sex-role implications regarding occupation.) For anyone slightly more sophisticated, however, there is too much which goes unsaid. For instance, how does one work with instructional goals in the affective domain? How are these specified? Where does one go from there once one has identified one's goals? Also, doesn't one run the risk of leaving out some important goals that one

really has? While additional books in this vein may help the teacher with his or her measurement problems, the affective area has been largely ignored because of the difficulties of clearly stating the behaviors one would like to see occur in students. The Fortune-Hutchinson methodology (Benedict, 1973a) handles this by means of the technique known as "The Operationalization of Fuzzy Concepts" (Coffing, Hutchinson, Thomann, and Allan, 1971), and the problem of incomplete goals lists by tests of completeness in the goals process and in several other of the major steps.

VIMCET associates of Los Angeles, with the help of Popham and Baker, have put out a series of thirty slide-tape modules dealing with such evaluation topics as identifying objectives, developing tests, analyzing learning outcomes, individualizing instruction, and systematic instructional decision-making. In "Humanizing Educational Objectives," by Baker (1972), the teacher is directed to stimulate students to appreciate the process of education, rather than merely its products, i.e., grades. Goals should reflect a concern for the individual, and students can be directly involved in a needs assessment prior to the selection of the goals. While impractical for very young children, older students can actually help select the objectives and develop the priorities on the basis of such criteria as interest, unmastered objectives, and importance.

In "Alternative Measurement Tactics for Educational Evaluation" by Popham (1972) for VIMCET Associates, the teacher is instructed in the use of criterion measures and other non-standardized devices. The teacher decides whether it is a learner product or a learner behavior in which (s)he is interested and whether the conditions will be natural or manipulated. Many examples of non-traditional measures are given.

As in the books on instructional objectives mentioned above, the slide-tape modules can prove very valuable to classroom teachers or anyone who needs to gain an initial understanding of the processes of goal setting and collection and analysis of data. This set of modules has an advantage, in this author's opinion, of offering something to the individual who would like to specify some affective goals or use instruments other than standardized tests. For conducting a complete evaluation, however, these slide-tapes would not provide a continual flow of steps or be operational enough to help remedy individual problems. The presentation mode holds the interest, and, if the modules could be more complete and represent a more systematic set of procedures, this could be a valuable tool for training beginners in educational evaluation.

Several other materials have been developed for training users to carry out a part or parts of the evaluation process. The Center for the Study of Evaluation (CSE) of

the School of Education of UCLA has developed the Elementary School Evaluation Kit. This appears somewhat cumbersome and difficult to use. It contains, for instance, ten packs of 106 goal cards each so that the principal and his or her community can help decide on the specific goals to be addressed by a particular school for a particular year. This system, however, appears to rely heavily on standardized tests, rather than other observational techniques, for gathering data concerning the goals, most of which are in the cognitive domain. CSE also developed a format for a workshop concerned with teacher appraisal, the emphasis here being on educational improvement, rather than merely on passing judgment on teachers. Instruments to measure and classify relationships between student-teacher roles and classroom interaction are suggested. This has the drawback, however, of ignoring the goals of individual teachers for their own classroom behaviors in the process of appraising these behaviors by a pre-determined instrument.

In its training program in educational development, dissemination, and evaluation, the Far West Laboratory for Educational Research and Development uses case studies as an introduction to various aspects of evaluation. The program, in a module format including student workbooks, includes such topics as: the uses of tests; norm vs. criterion-referenced tests; critical properties of evaluation instruments; sources of evaluation instruments, and how to select

and evaluate them; and a section on evaluation problems. While of potential use to the novice evaluator, this program is not prescriptive enough to be of real value to someone who wants to carry out a complete evaluation for the first time. Even with these materials, the practitioner would be left to his or her own devices to fill in the many gaps between steps or sections.

Rather than continuing attempts to arrive at a complete evaluation training program, the authors of two different studies started with a search for the evaluation skills needed to be taught by such a program. (Note: The intention of this author is to cite studies concerning evaluation training in which evaluation is defined to be the providing of data for decision making (Cronbach, 1963) which is the definition employed in the Fortune-Hutchinson methodology and the one used by Guba and Gephart (1970). In addition, Worthen, Anderson, and Byers (1971) have been cited because of their recognition of evaluation as a legitimate activity in its own right, one which is separable from disciplined inquiry and which is not solely equated with measurement.) In "Training Materials for Research, Development and Diffusion Training Programs," Guba and Gephart (1970) identified several such skills, and, for the same study, Eboch and Stufflebeam worked on developing the instruction to teach them. Because the skills were not stated specifically enough, however, the resulting twenty-two overhead transparencies

with coordinated twenty-two page script have the same shortcomings of most of the materials mentioned above.

In "A Study of Selected Factors Related to the Training of Researchers, Developers, Diffusers, and Evaluators in Education," Worthen, Anderson, and Byers (1971) interviewed a random sample of educational researchers engaged in exemplary educational research and research-related work. From a factor analysis of 116 interviews, twelve "task factors" were isolated. These included such things as: design of research studies; conducting and interpreting data analyses; developing instructional material; conducting evaluations; and constructing and using data collection instruments. In addition, 226 competencies, including both skills and knowledge, were first judged against three criteria and then tabulated under each task category. The criteria were: that it be trainable by means of a systematic training technique, rather than a personal characteristic; that it not be the end product of training in some other field; and that it be important or significant, i.e., not trivial.

Two forms of a test of educational research and evaluation competencies were developed. They seemed to be more valuable for the research than for the evaluation area, however, because the evaluation competencies were not operationally stated and most pertained only to a knowledge of the literature of evaluation. An example was given for clarification in each instance, but often some ambiguity still

remained. This skill-identification approach, of all those mentioned, however, seems to this author to be the most valuable for developing a sound, systematic, replicable training program for evaluators.

The Methodological Approach

As described above, there are several problems with existing evaluation training materials. The most notable of these problems is that all lack completeness. None of the approaches to evaluator training offers a full, systematic set of steps for carrying out an evaluation. A major reason for this lack could be that none of the training materials apparently was developed with a defined purpose in mind.

In his "A Critique of Federal Evaluative Methodology," Hutchinson (1972a) compared the model used by federal agencies for conducting their evaluations with the Fortune-Hutchinson Evaluation Methodology. The former lacks a defined purpose and criteria for judging the success of the evaluations, conditions which may lead to the collection of data which aren't subsequently used to make decisions, thus wasting the resources expended. This assumes that the desired aim of data collection is the making of informed decisions for the purpose of improving a program. Without a defined purpose, however, one really does not even know why data are being collected. The data gathered in federal evaluations, according to Hutchinson (1972a), do not relate to the goals of decision makers, although they may relate instead to the

goals of those who designed the evaluation but who do not necessarily have to make decisions about the programs.

In comparison, the Fortune-Hutchinson Evaluation Methodology is a systematic, operational, standardized set of rules and procedures to provide data for decision making (Hutchinson, 1972a) which also contains criteria for determining its own success or failure. This methodology preserves decision maker validity throughout, i.e., it guarantees that the data provided are those which the decision maker really wants and will use.

A systematic set of steps also has the advantage over the haphazard approach in that it can continually be improved upon because one always knows the procedure being employed. Thus, methodological development can proceed through monitored field testing of the methodology in the hope of producing more successful evaluations (Hutchinson, 1972a).

For the reasons listed above, this author is focusing her efforts for producing improved evaluation training materials within a methodological framework. That framework is the Fortune-Hutchinson Evaluation Methodology which will be more fully described in the next section.

The Fortune-Hutchinson Evaluation Methodology

Since the concern of this investigation is the development of training materials for evaluators using the Fortune-Hutchinson Evaluation Methodology (hereafter to be referred to in this text as "F-H"), it is appropriate to provide some

background on its development and to summarize its content.

The nature of F-H. F-H is a methodology designed to provide data for decision making. As Gordon (1973) has shown, this purpose is desirable, practical, and operationalizable, and other existing methodologies are insufficient to accomplish this purpose. The most important implication of this purpose is that the evaluation should provide data that the decision maker actually used, i.e., data which have decision maker validity. In order to insure that this happens, the decision maker's goals must be fully identified and fully specified. Where resource limitations are present, the most important goals should be specified.

Benedict (1973c) has listed four assumptions behind the use of such a methodology. These are as follows: the decision makers of the enterprise or project have the moral and ethical right to make their own decisions about the enterprise; it is the responsibility of the decision makers of the project or enterprise to make their own decisions, not that of an outside "expert" or "consultant"; the only legitimate purpose of educational evaluation is to provide information to these decision makers for their own use as they see fit; and that the validity of this approach is ultimately determined by whether and to how great a degree the data are used by the decision makers in making decisions about the enterprise.

Finally, given this defined purpose and these assumptions, Hutchinson (1972c) inferred three criteria for a successful F-H evaluation. First, the evaluation should be as

efficient as possible. That is, the amount of data provided by the evaluation which is actually used by the decision maker in making decisions should be maximized. Next, of those decisions made by the decision maker, the number made using evaluation-provided data should be maximized. This is the criterion of completeness. The third criterion is that of focus which states that when the evaluation lacks 100% completeness, data should be provided for the important, rather than the unimportant, decisions which the decision maker has to make.

Major steps of F-H. The major steps of F-H and brief descriptions of their content are listed below.

1.0 Negotiation of the contract.

1.1 The individual in charge of the evaluation resources (the contract decision maker) is identified. The evaluator and contract decision maker determine whether or not F-H will be suitable for the evaluation of the enterprise in question.

1.2 The scope of the enterprise to be evaluated is determined.

1.3 The resources for the evaluation are determined by the evaluator and the contract decision maker. Others are then called in to do the same thing, thereby testing the completeness of the original list of resources.

1.4 Decision makers for whom data are to be collected are identified by the contract decision

maker and others. The final list is put in priority order, using one or more agreed-upon criteria and a reasonable number of decision makers, given the identified resources, and selected.

1.5 The contract is drawn up including the above information, and the final approval and signature of the contract decision maker is secured.

2.0 The contract decision maker reporting process.

2.1 The contract decision maker chooses the kind of reporting process (s)he would like to have implemented, given the information provided by the evaluator concerning possible choices for the content and frequency of the projected evaluation reports.

3.0 Allocation of resources among the parts of F-H.

3.1 The evaluator secures the cooperation of the decision makers, explains the evaluation contract to them, and asks them to keep a log of the decisions they make and the data used with which the decisions were made. For each decision maker, resource allocation charts are drawn up, and his or her portion of the evaluation resources are allocated to the various steps of the methodology, using the percentages suggested by the evaluator.

- 4.0 The goals process. (Cycle through once for each decision maker.)
- 4.1 The first (or next) priority decision maker is identified, and it is determined whether the decision maker is an individual or a group. If the decision maker is a group, it is determined how large the group is with respect to available resources and whether the group makes individual decisions or acts as a single decision making body. Separate procedures are included for each kind of decision maker. They are described below for an individual.
- 4.2 The decision maker responds to an open-ended question such as "What do you want or intend (the enterprise)* to be and to accomplish?"
*The evaluator substitutes the name of the enterprise.
- 4.3 The evaluator performs a goal analysis on the responses.
- 4.4 Alternative goals lists are developed using selected enterprise documents and the responses of other decision makers to the above stimulus question.
- 4.5 The decision maker reacts to these lists by adding the goals to his or her own list,

changing any goals (s)he wishes to change or adding any new goals (s)he happens to think of.

- 4.6 If resources allow, an activities test of completeness is performed.
 - 4.7 The decision maker commits himself to his goals list.
 - 4.8 The decision maker puts his or her goals in priority order using one or more criteria previously decided upon with the help of the evaluator.
- 5.0 The parts process. (Cycle through once for each decision maker.)
- 5.1 Separate procedures are also used here depending on the kind of decision maker (group or individual) for whom the process is being followed.
 - 5.2 The decision maker responds to an open-ended stimulus question such as "What are the conceptual components that you see as the major parts of (the enterprise)*?" *The evaluator substitutes the name of the enterprise.
 - 5.3 The decision maker is asked to categorize the parts identified above as "Inputs," "Interfaces," and "Outputs" where these are defined in the following way.

Inputs are those things occurring before the enterprise begins or those prerequisites for the program, e.g. in a school situation these might be budget, physical plant, etc.

Interfaces are those things which are not directly part of, but which impinge upon, the enterprise and thus influence it, e.g. School Board, PTA, etc.

Outputs are those things resulting from the project or program occurring after the program is ended. For example, an output in a school situation might be the student at the end of the program or school year.

The decision maker is asked to consider whether each category is as complete as possible. If not, the necessary parts are added to the original list.

5.4 Others respond to the open-ended stimulus question and the decision maker reacts to their lists, as in the goals process.

5.5 If the activities test of completeness was performed for goals, that list is used for the decision maker to match parts to activities.

5.6 Where needed, subparts of parts are identified using the above process.

- 5.7 Parts (and subparts) are put in priority order, as were goals.
- 5.8 The evaluator obtains decision maker commitment for the list of parts.
- 6.0 Integration of goals and parts. (Cycle through once for each decision maker.)
 - 6.1 The "Goals, Activities, and Parts Matrix" developed by Stetz (1972) is used to match goals to parts and vice versa.
- 7.0 Operationalization of goals. (Cycle through once for each goal of each decision maker.)
 - 7.1 Different procedures are used depending on whether the decision maker is an individual or a group.
 - 7.2 The evaluator takes the decision maker through "The Operationalization of Fuzzy Concepts" (Hutchinson and Benedict, 1970a) using the highest (or next highest) priority goal used in the context of the identified goal-part interface. This involves conceptualizing and stating a hypothetical situation in which the goal is optimally being achieved. The decision maker is asked to state or write down all the things going on in the hypothetical situation which indicates that the goal is being achieved in the best possible way.

The decision maker is also asked to respond to a similar situation except that the goal in this case is not being achieved at all. Finally, three tests of completeness are done. First, the decision maker responds to lists obtained from others who have responded to the two hypothetical situations. Secondly, the decision maker carefully reexamines the hypothetical situations, searching out and considering anything that might have been left out. Lastly, the decision maker thinks of things which have nothing to do with the goal in question and then seriously examines whether or not they do.

- 7.3 Everything written down on the decision maker's list is considered a "goal component" and are all put in priority order in terms of the importance of having evaluation data about them.
- 7.4 Items on the list which are not directly observable are put into priority order in terms of the above criterion.
- 7.5 The operationalization process continues with the most (or next most) important goal-component which is not directly observable.

8.0 Design of measurement techniques. (Cycle through once for each goal-component of each decision maker.)

8.1 If the evaluator cannot perform the steps in this section, a measurement consultant is necessary.

8.2 The ideal measurement technique is designed, i.e. it is planned how the goal component can be measured directly, unobtrusively, and under natural conditions.

8.3 Taking available resources into account, a measurement technique as close to the ideal as possible is designed.

8.4 If resources allow, the technique should be field tested and tested for validity with all problems noted.

8.5 The decision maker approves the technique or requests that it be redesigned. If a redesign is requested, it will be carried out using the above procedures and the decision maker's response.

9.0 Implementation of measurement techniques.

9.1 It is determined whether or not sampling is required.

9.2 A recording device is developed and field tested.

9.3 If needed, a sampling plan is developed and tested for completeness.

9.4 The observations are carried out.

9.5 It is planned when the observations will be repeated according to the above procedures.

10.0 Reporting

10.1 There are different procedures depending on the type of decision maker to whom data are to be reported.

10.2 The report is written including vital information about the goal component, its priority order, and its related parts.

10.3 The data are presented clearly, preferably in a chart, table, or graph where appropriate.

10.4 The data are given to the decision maker who is asked to read it.

10.5 Difficulties in interpretation of the results are pointed out by the evaluator.

11.0 Redesign.

11.1 A report on all evaluation activities performed to date for the decision maker requesting the redesign is prepared by the evaluator.

11.2 If the evaluator rather than a decision maker initiates the redesign, these procedures commence with the highest priority decision

maker for whom the evaluator wishes to redesign the evaluation.

11.3 If the contract decision maker requests the redesign (or in the case of normal redesign for a long-term evaluation), the evaluator prepares a report on all evaluation activities performed to date.

11.4 The design procedures that are being redone are gone through again using the original output of the decision maker as the only test of completeness.

12.0 Evaluation of the evaluation.

12.1 The evaluator, using the decision maker's log of decisions and data, determines the extent to which evaluation-provided data have been used for decision making.

12.2 The evaluator calculates percentages of completeness, efficiency, and focus.

12.3 Under certain conditions, these percentages may indicate to the evaluator that a redesign of the evaluation is needed.

The Importance of the Fortune-Hutchinson Evaluation Methodology

The potential value of F-H has been well-documented to date (Gordon, 1973; Benedict, 1973b; Rosen, 1974). In this

author's opinion, its biggest advantages lie in the fact that it is a methodology with the sole purpose of providing data for decision making and that it contains criteria for the assessment of the degree of accomplishment of that purpose--the criteria of completeness, efficiency, and focus previously mentioned.

The need for evaluation methodology cannot be filled by turning to the area of educational research. Guba and Stufflebeam (1968) noted that in an evaluation there is a need for continual, rather than post hoc, data. The aim is for specification of the degree of accomplishment of the goals of a particular program, rather than the generalizing of results, which is one of the aims of good research design. In a research experiment, the treatment can't be modified during the course of the experiment, while in many evaluations, continual data feed-back hopefully will lead to ongoing improvements in the program. An evaluation is usually concerned with a whole program involving many interacting variables, while the research ideal is better suited to dealing with a single unit with an easily-controlled number of variables. An evaluation also can rarely include random assignment of students to treatments, as ethical considerations often do not permit the withholding of a special program from a designated "control" group. The notion of random assignment is vital to research, however, where the focus is on generating generalizable knowledge.

Thus, F-H is a real answer to the need for evaluation methodology in the field of education. While not attempting to duplicate procedures from the field of research, it still provides a systematic, standardized, operationalized set of rules and procedures for providing data for decision making.

In addition to having a purpose different from that of traditional research methodology, F-H has been found to have advantages for use in alternative schools for which traditional evaluation approaches have short-comings. In his "Evaluation For Alternative Schools," Rosen (1972) defines alternative schools in a general way as schools which have very different goals, values, and styles from the common neighborhood public school, i.e. schools which have as their basic commonality the fact that they differ significantly from traditional schools.

For this kind of school, it is important not to limit measurement instruments to those which have been shown to be valid and reliable as has been the tendency of many evaluators who continue to equate evaluation and testing. The reason for this according to Rosen (1972), is that no instruments are presently available to measure some of the most important goals of alternative schools, such as affective goals. In F-H, however, goals are operationalized solely in terms of the decision maker's meaning for the goal. Thus, the resulting measurement technique, while it may not

be reliable over time, is as valid as it can possibly be for the decision maker in question.

Thus, F-H is capable of supplying data for alternative schools which are really desired by their decision makers and which directly pertain to their goals. Rosen goes on to mention that F-H has an advantage over other evaluation models in that it is a systematic, standardized, operational set of rules and procedures which can be learned and carried out by someone from inside the school, thereby eliminating the necessity of hiring an outside evaluator. This may be a very important consideration, especially when funds for evaluation are small.

The Training of Fortune-Hutchinson Evaluators

At present, this evaluation methodology is largely taught by means of copies of sections of the methodology itself which are first read by the students and later discussed and explained in class. Because Hutchinson believes that any methodology should be subject to continual development, the sections of the methodology are constantly being updated as gaps in the logic and/or workability are filled, particularly through the results of field testing. Thus, students of evaluation are presented with a complete, up-to-date version of the steps from which to learn the complete process. These materials have been combined by Gorth, O'Reilly, and Pinsky (1973).

Four of the sections of the methodology, the goals process, the goal analysis procedures, the goals-parts interface process, and the operationalization process, have been put into self-instructional workbook format by Benedict, Rosen, and Hutchinson (1973); Benedict (1973b); Thomann (1972); and Coffing, Hutchinson, Thomann, and Allan (1971), respectively. While the original aim was to facilitate the work of the evaluator by providing the workbooks to decision makers, rather than using them to train evaluators, the workbooks could be used for the latter purpose, either as a check on the effects of instruction, or as a pre-instructional tool.

A graduate student working in the Student Affairs Research and Evaluation Office of the University of Massachusetts at Amherst, has taken the Fortune-Hutchinson Evaluation Methodology and created a series of handbooks for use in training personnel in evaluation to help them carry out their management planning (Brooks, 1974). The identified functions differ slightly from the sequence of steps in the original methodology, but the purpose and rationale of each function are included, as well as its relation to the whole and an example for clarification. The presentations are clear, both in language and format, and an overview of the entire process is presented in the first handbook.

Other existing training materials for this evaluation methodology include an instructional system developed by

this author, workshop materials developed by Hutchinson and Benedict (1970b), and a course in evaluation taught at Southern Connecticut State College (Rosen, 1974). The instructional system included needs and task analyses of the content of the proposed methodology course, along with behavioral objectives and criterion-referenced test items. The workshop materials contained training and workshop experiences in some of the skills important in using the methodology, e.g., goal identification and operationalization. Finally, Rosen's materials include a complete set of goals the instructor had for the students' achievement, criteria for student success, and a course evaluation to be filled out by the students.

Statement of the Problem

The preceding analysis of the present state of evaluation training materials in general, and F-H training materials in particular, indicates the lack of a systematic way of training F-H evaluators. This can lead to insufficiently trained evaluators who are forced to learn "on-the-job" to cope with problems that arise, a situation that can be frustrating for them and the decision makers with whom they are working.

General Description of Procedures

Because of the need for improved training materials for F-H evaluators, it is this author's opinion that the skill identification approach (Worthen, Anderson and Byers, 1971)

will be used here. The skills will be put in order of importance, and training materials developed and field tested (Tyler, 1950) for as many of the top skills as can be reasonably treated within the time allotted for these activities.

Some General Principles for Developing and Evaluating Training Materials

One of the clearest articulations of basic principles for the development of curricula in general was put forth by Tyler in his Basic Principles of Curriculum and Instruction (1950). He considered education to be a process of changing behavior patterns in people, where behavior in this case includes thinking and feeling, as well as overt action. The educational objectives then represent the kinds of changes in behavior which one seeks to bring about in students (p.4).

Tyler discussed several possible sources for these objectives. The students' needs should be studied so that those needs being met by the home and community wouldn't be duplicated in the schools which would concentrate, instead, on the gaps in student need fulfillment. Student interests should be considered so that educational objectives may foster active student participation in the learning process. Other sources could include studies of contemporary life, the suggestions of subject matter specialists, and the field of psychology of learning.

Once the objectives are identified, learning experiences must be designed to meet the objectives because, in Tyler's view, it is the experiences provided, rather than merely the

things to which the student is exposed, which are essential to education (p. 41). Among his general principles for planning learning objectives, Tyler stressed that students must be able to practice the behavior implied in the objective and its content area. Also, a learning experience must be such that a student gains satisfaction from carrying on the behavior. (Skinner, who is discussed below, would say that the student must be reinforced in order for learning to take place.) The reactions desired in the learning experience must be within the range of possibilities for the students involved, given their present attainments and predispositions. Finally, different learning experiences can be used to attain the same educational objectives, and the same learning experience may often bring about several different outcomes for different people (pp. 42-44).

Tyler emphasized that educational changes take time. The curriculum must be unified by one or more organizing elements, e.g. concepts, values, or skills, with the learning experiences organized so that they reinforce one another. Organizing principles for the learning experiences within the curriculum may be chosen from among the following: chronological order; increasing breadth of application; increasing range of activities included; description followed by broader and broader principles to explain these illustrations; and an attempt to build an increasingly unified world picture from specific parts which are first built

into larger and larger wholes (pp. 63-64). Effective organization, in Tyler's view, is judged by three criteria. The criterion of continuity involves the reiteration of the major curricular elements, or vertical organization. The criterion of sequence requires that each successive learning experience build upon the previous one but go more deeply into the topic. Finally, the criterion of integration implies the horizontal relationship of curricular experiences (p. 55).

In many respects, Skinner and Tyler are in agreement with respect to principles for planning curriculum, even though the former came from the realm of the laboratory and research into animal behavior. For instance, Skinner wrote in his Technology of Teaching (1968) that "the first step in designing instruction is to define the terminal behavior" (p. 199). He stresses, however, that teaching is the ". . . the arrangement of contingencies of reinforcement," i.e. ordering the relations which prevail between behavior on the one hand and the consequences of that behavior on the other (p. 5).

There are many things that might reinforce a behavior. For example, the student's merely being able to exercise control over nature by mastering a skill serves to reinforce behaviors leading to the mastery of the skill. In addition, being able to do what one wants with a block of time, to compete with others (which competition is only reinforcing

when one is successful), to earn the good will of the teacher, and to avoid aversive stimulation all can reinforce desirable behavior. As Tyler (1950) put it, the student should gain satisfaction from carrying on the behavior.

Reinforcements must be made contingent upon the desired behavior. In the case of very complex behaviors, they must be divided into large numbers of very small steps, so that the reinforcement can then be contingent upon the accomplishment of each step. By making each step as small as possible, the frequency of reinforcement is raised to a maximum, while the possibly aversive consequences of being wrong are reduced to a minimum. The material must be made easy to remember and be well organized.

Skinner (1968) went on to explain that all learning results from the initial behaviors on the part of the learner. Tyler also stressed that the learning objectives must be within the range of attainment, given the students' present abilities. Skinner wrote that the teacher must induce the behavior so that it can be reinforced, but must be concerned about the probability for the learner repeating the behavior outside of the school environment. Thus, the teacher must help the student become free of the instructional contingencies and respond eventually to the contingencies of the world at large. To do this, the teacher must know the strength of the student's entering knowledge so

that the ratio of reinforcement to activity can be gradually stretched out.

One concept which Skinner brought up which wasn't touched upon by Tyler was Skinner's idea that, as a reinforcing mechanism, the teacher is out-of-date, even if his or her entire attention could be directed toward a single child! Since a lapse of only a few seconds between response and reinforcement can destroy most of the effect of the latter, especially for mathematical behaviors, a lapse of as much as twenty-four hours, which is not uncommon in many classrooms, renders the process untenable. This delay results from the teacher not being able to reinforce each step in a series of progressive approximations to final complex behaviors because (s)he can't deal with the pupil's responses one at a time. Too many students in a classroom and too many steps before mastery of the behavior make proper reinforcement impossible. Skinner concluded that the most effective control of human learning will require instrumental aid. One such instrument, programmed instruction, will be discussed in a later chapter.

With respect to learning theory, Gagné (1973) is in agreement with Skinner. That is, he believes that learning takes place when the instructional stimulus causes a response in the learner which is then in some way reinforced. Possible stimuli are of several different kinds, but unless they are relevant to the task at hand, the instructional

sequence won't be effective. The instructional sequence, defined by Gagné as a succession of events with the purpose of insuring that the necessary stimulus conditions for a single learning act will all be present at the proper time, requires that those elements that need to be present in the immediate situation be distinguished from those which must be retrieved from memory (p. 9). In addition, the instructional sequence should be designed " . . . in relation to the type of objective for the learning as a whole and this depends, of course, on how the learning is to be measured" (p. 10, underlining is the author's).

Merrill and Boutwell (1973) stressed the difficulty of specifying learning behaviors. Content of the task, as well as desired behavior, must be taken into account. Qualitative variables, such as presentation form, inter-display relations, and mathemagenic information (additional information facilitating learning, e.g. prompting and feedback), and quantitative variables such as sequence, quantity, and pace contribute to the quality of the instructional sequence being presented.

In addition, according to de Cecco (1968), the students' entering behavior and motivation must be considered. De Cecco enumerated several variables constituting entering behavior. Among these were training, maturation, individual differences, and personality. This parallels Tyler's discussion of the necessity for examining student interests and

needs before planning instructional objectives. De Cecco also described several functions which the teacher should undertake in order to increase student motivation. These include: engaging the student in learning; describing instructional objectives in advance; reinforcing achievement; and using rewards and punishments to control deviant behavior.

While not intended to be an exhaustive treatment of current approaches to curriculum development, the above discussion concerns common principles which were used by this author to develop instructional materials for F-H evaluators. Some more recent writing in the curriculum field have tackled broader problems than those discussed here, such as the need for working theories of curriculum development (Schwab, 1969); the need to focus on the enduring problems of education (Schwab, 1969); the need for an examination of criteria used to make decisions about the elements of a curriculum (Kirst and Walker, 1971); an increased stress on curricula that demonstrate interrelationships across disciplines (Goodlad, Von Stephasias, and Klein, 1966; Schwab, 1969); and the need for more curricular experimentation (Goodlad, Von Stephasias, and Klein, 1966). However, all support the general guidelines summarized below and first put forth in Tyler's Basic Principles of Curriculum and Instruction (1950) which has been called "The primer for curriculum planners in any educational enterprise; specifies some basic questions

and suggests where some of the answers may be found" (Goodlad, Von Stephasias and Klein, 1966, p. 119).

The authors above agree in general about how curricula are most effectively planned. Their approach includes doing a needs analysis for the students in the content area for which instruction is contemplated. Then, taking motivation into account, specific behavior changes to be brought about by the instruction are carefully specified, and learning experiences are designed to provide necessary stimuli and reinforcements to produce those changes.

Evaluating effectiveness of learning experiences. As interpreted by Tyler (1950), this evaluation is a process for finding out how far the learning experiences as developed and organized are actually producing the desired results, and for identifying the strengths and weaknesses of the plans. This implies that the original student behavior must be appraised and that there be more than one appraisal thereafter to determine what changes are occurring, when these changes occur, and the permanence of the changes.

Tyler listed several types of assessment techniques which can be used in addition to pencil-and-paper tests which are most appropriate for testing writing and verbal abilities. Some of these are: direct observation for assessing habits, skills, and personal-social adjustment; interviews for determining attitudes, interests, and appreciation; questionnaires also for examining interests and

attitudes; the collection of various products such as themes, paintings, and shop materials; and the examination of records such as health records, numbers of library books withdrawn, etc.

For the planning of the evaluation, Tyler suggested that one start with the procedures for planning objectives in order to see which behaviors should be appraised and which content areas should be sampled. It is important to define situations which will give students opportunities to express the behaviors implied by the educational objectives, i.e., situations sufficiently under control to be evaluated. When objectives and situations are identified, available instruments should be reviewed to see if any might serve the purpose.

In any case, the instrument should be tried out before using it for formal collection of data in order to determine how observations are to be recorded and the terms of units to be used to summarize the data. This summary will preferably be analytic, rather than a single score. It is necessary also to determine to what degree the instrument is free from observer bias or error, how adequate is the sample of behavior included in the instrument, how frequently the behavior is assessed, and whether the instrument is long enough to fully evaluate the behaviors. One must also decide whether the instrument either samples directly or correlates highly with direct evidence of the behaviors in question. This is a question of validity which De Cecco (1968) defined as the

test measuring what it purports to measure. Tyler went on to say that the results of test administrations should be examined before and after given periods in order to estimate the amount of behavior change taking place. Similarly, this is a question of reliability which De Cecco defined as the consistent measurement of terminal performances over repeated administrations. This indicates that items are neither too easy nor too difficult, the items aren't ambiguous, the scoring is objective, and the student behavior being measured is consistent.

Finally, Tyler (1950) concluded that one must form and check out hypotheses about the results of tests and observations. The curriculum should be modified accordingly and checked out again by re-teaching the material and subsequently re-evaluating it.

Thus, the general principles for developing and evaluating instructional materials are as follows:

1. Do a needs analysis of the audience for whom the instruction is intended.
2. Specify behavior changes to be brought about by the instruction, considering identified needs.
3. Take motivation and entering behaviors into account.
4. Design learning experiences which will bring about the behavior changes and will also provide reinforcement for these accomplishments.

5. Design tests to determine the accomplishment of the behavior changes and pre-tests to determine initial knowledge, by defining situations in which students can practice the behaviors.
6. Field test the instructional materials and tests and revise them before large-scale use.
7. Administer the test repeatedly to determine its reliability, and determine its validity by investigating the degree to which it correlates with direct measurement of the desired behaviors.

Skill Identification: The First Step

With the summary of general principles for developing curricula in mind, it is evident that the development of training materials for F-H is lacking in the area of a clear statement of behavior changes intended to be brought about by the instruction. It has not yet been clearly stated what one needs to be able to do in order to actually be an F-H evaluator. Of course, one needs to understand and be able to use the methodology, and this is the present emphasis of F-H courses. This is, however, a very broad objective. Are there skills that one must have before attempting to learn F-H? Are there skills which need to be developed while learning F-H but which are not specifically taught by this methodology? It is necessary to answer these questions

before attempting to specify objectives on which to base improved F-H training materials.

Outside the F-H framework, as stated above, there have been efforts made to specify skills required by evaluators in general. The study by Guba and Gephart (1970) began by establishing several criteria for a good evaluation. These included internal validity, external validity, reliability, objectivity, relevance, significance, scope, credibility, timeliness, pervasiveness, and efficiency (p. 25). They commented that existing evaluation training programs tend to handle the first four well but to miss on the last seven. Evaluators, they felt, should be able to do the following:

1. Focus evaluation efforts so that the alternatives possible for a specific decision are compared on criteria which decision makers will use in arriving at a choice.
2. Collect information on each decision maker's criterion for each alternative.
3. Organize information into a format understandable by and acceptable to the decision makers in question.
4. Analyze and interpret the information.
5. Report the information to decision makers.

In addition, Guba and Gephart included skills in interpersonal relationships, small-group dynamics, and systems analysis.

Their resulting materials based on this rough skill identification included: a description of terms and concepts inherent in the evaluation process; a definition of evaluation; lists of categories of decisions; and roles inherent in evaluation and decision making. As previously noted, however, the results of the skill identification were too general to lead directly to a statement of specific behavioral objectives. Thus, it is difficult to say whether or not the resulting training materials really do what they should be doing.

The Worthen, Anderson, and Byers study (1971) discussed above seems to this author to indicate the most productive line of inquiry in order to produce effective, accountable training materials for F-H. A thorough skill identification process and a break-down of identified skills where appropriate to eliminate ambiguity is necessary for a clear statement of instructional objectives which, in turn, is required for the development of sound instructional materials and instruments for the evaluation of instruction.

CHAPTER II

GENERAL DESIGN OF THE STUDY AND THE SPECIFIC
DESIGN OF THE SKILL IDENTIFICATION PROCESSGeneral Design of the Study

From the statement of the problem and the specific curricular guidelines given in Chapter I, this study was conceived as consisting of three parts. These are: the skill identification procedures which include a thorough search of many possible sources of skills for F-H and putting this list in priority order by skill importance; the development of training materials and tests for one or more of the top skills, requiring a statement of instructional objectives and possibly the further clarification of the skills to be used; the field testing of the materials so that major problems may be identified and revisions planned. These procedures are necessary to insure a skills list which will be optimally helpful to future trainers of F-H evaluators and to provide an example of how it may be used to produce appropriate instructional materials.

The skill-identification approach is one means of doing the student needs analysis advocated by Tyler as a prerequisite to the development of meaningful instruction. In this way the needs of future F-H evaluators for specific skills are identified, put into order with respect to importance for successfully using F-H, and then, if still unclear, are specified more fully.

The development of instruction and the means for evaluating this instruction can then come directly from the statement of the specified skill which is the source of specific educational objectives. The field testing of the resulting materials, also discussed by Tyler, is a means of improving instruction even before widespread administration has occurred. This process should, of course, be a continual feed-back, allowing for ongoing improvement of the materials.

Specific Design of Skill Identification Procedures

The design of techniques which were used to identify the skills needed for an individual to become a successful F-H evaluator began with a consideration of all available sources. When several possible sources had been identified, they were put in order according to their apparent importance, with the ones on the top of the list commanding a more thorough effort and a greater expenditure of time than those lower on the list. It was decided that, since this was the first attempt to identify skills required by F-H and was potentially of great value for future curriculum development in this area, that the emphasis would be on quantity of skills identified, rather than quality. Thus, this author was not terribly concerned if some skills appeared vague or if skills overlapped. Also, there was no distinction made between "skills" and "knowledge." It would have been time-consuming to completely specify each skill, leaving no time for the development of training materials. Therefore, the

skills list was left alone which resulted in some problems which will be discussed later in this chapter.

Search of F-H documentation. The most up-to-date documented form of the Fortune-Hutchinson Evaluation Methodology written by Hutchinson, et al. and compiled by Benedict (1973) seemed to be the logical starting point for the skill identification process. It was assumed by this author that an in-depth examination of the documentation would yield many steps which assumed skills or knowledge on the part of the reader. Two approaches were used to elicit required skills from this form of the methodology. For the first, this author went through the documentation in a step-wise fashion, and for each step and sub-step asked the question "Does one need to know anything or be able to do anything in order to accomplish this?" If the answer was "yes," the required knowledge or skill was put on a list. When the end of each major step was reached, this author went through it again, this time forming small clusters of related sub-steps and asking the same question. For the second approach, the clustering was put before the examination of the individual sub-steps. The clustering was not done in a systematic way, but each clustered group of steps was as large as necessary to give this author a more general interpretation of the procedures being discussed. The purpose of using both methods was to see whether one or the other yielded more skills for the list. The first approach yielded more skills and was

therefore used for the balance of the documentation after the first major step.

User questionnaires. The second most important method used to identify skills needed for F-H was a brief questionnaire distributed to students of the methodology. Nine graduate students responded, all of whom had done evaluations using F-H and were familiar with the methodology. The questions were as follows:

1. What skills did you have that helped you to do the evaluation task?
2. What skills do you feel you developed by learning the methodology?
3. What skills do you wish you had that would have made your use of the methodology more successful?

There was some confusion on the second question, and at least one person interpreted it to mean skills developed by applying the methodology in a real school context. The process was, however, quite productive and yielded many skills that had not been identified by means of the search of the documentation. All responses to all questions, except where completely redundant, were added to the first list.

Search of related dissertations. The third source of skills needed by an F-H evaluator was dissertations relating to the methodology. Five dissertations were used, three of which involved field tests of a major step in F-H (the goals process, Benedict, 1973; the operationalization process,

Jones, 1970; and the measurement process, Jeffers, 1974). The other two were field tests of the methodology as a whole (Gordon, 1973, and Rosen, 1974). Chapters entitled "Recommendations," "Directions for Future Research," and the like were carefully reviewed for any mention of a need for improved training materials for evaluators or for methodological gaps with respect to evaluator training. Some skills were identified in this manner, but not as many as by either of the first two techniques. New skills were added to the above list.

Search of other evaluation models. It did not seem to make much sense to do an in-depth search of other evaluation models for three reasons. The first is that F-H is a methodology which emphasizes the decision maker role in evaluation. Because it is a methodology, it is much more operational than a decision maker model could be (Hutchinson, 1972c), and, therefore, it seemed unlikely that a model could provide any but very general skills pertaining to F-H. A model offers suggested procedures, as opposed to the systematic, standardized, operational ones given in F-H. Thus, it would be very unlikely that any skills identified from a model would relate to any specific step in F-H. It would be more likely that skills obtained from models either would not be applicable to an F-H evaluation, or would end up at the bottom of the final list when it is ordered on the criterion "importance for being a successful F-H evaluator." The second reason is that

this dissertation involves the training of F-H evaluators, not evaluators using any other model or methodology, thus diminishing further the chances of finding skills appropriate for F-H from other models. Finally, by this time, the skills list was lengthy and apparently quite complete, and it didn't seem that a further search would be very productive.

In spite of these reasons, a few skills were added from outside sources. These sources included the general recommendations made by Stake and Denny (1969) and Worthen and Sanders (1973), and the two studies involving the identification of skills needed for evaluation (Guba and Gephart, 1970; and Worthen, Anderson, and Byers, 1971).

Putting skills list in priority order. Once as many skills as possible had been accumulated using the above techniques, it was necessary that they be put in some kind of order, since training materials, because of time constraints, could only be developed for a few skills at the top of the list. It was decided that the ordering be done on the basis of the criterion "Importance for being a successful F-H evaluator," and that the final list be a combined weighting of three separate ordered lists--one done by this author, one by Dr. Hutchinson of the University of Massachusetts at Amherst, co-originator of F-H, and one by Dr. Benedict, also from the University of Massachusetts, who has had a great deal of experience applying, teaching, and developing the methodology.

The method used was to group the resulting skills into six categories: knowledge of the methodology; knowledge from experience (use of the methodology); interpersonal skills; statistical and quantitative skills; resource skills; and other knowledge. Each of the three persons mentioned above rated each of the six categories from one (most important) to six (least important). Each skill was then placed on one of five rating levels, from one (most important) to five (least important). Each person's final list was derived by combining the importance level of the category with that of the skill itself, e.g., a skill from the most important category which had been placed on the first rating level would be placed on the highest level on the final list. A skill from the second most important category which had been placed on the first rating level would have the same final priority as a skill from the most important category which had been placed on the second rating level. Each final list, therefore, had thirty possible levels on it.

The combined list resulted from adding weights assigned to each individual list. Each level on this author's list was multiplied by a "1," on Dr. Hutchinson's list by "2," and by "3" on Dr. Benedict's. This was done to give an advantage to this author by pushing her most important skills closer to the top of the combined list. It was decided that this was appropriate because it was this author who would be developing the training materials and would be more likely to be

motivated for this task if she were working on skills she felt were most important. Dr. Hutchinson's list was placed above Dr. Benedict's because of his greater theoretical experience resulting from his development of F-H. Skills on the top of the combined list would be those which had the lowest total numerical priority attached to them. Thus, if the ordered position of each of this author's skills is noted by an "A," Dr. Hutchinson's by a "B," and Dr. Benedict's by a "C," the highest possible skill on the combined list ($A+2B+3C$) would result when $A=B=C=1$, or $A+2B+3C=6$. The range of values for skills' levels, then, is from 6 to 180, since there is a maximum of thirty levels on each individual list, and, when $A=B=C=30$, $A+2B+3C=180$.

Specifying skills on the list. After the list of skills was put in final priority order, it was clear that at least one of the skills for which instruction would be developed was not stated specifically enough. It was decided that the skill "The evaluator should be able to deal with a lack of decision maker cooperation" needed to be further specified by a discussion among professors and/or graduate students for whom the skill or skills had meaning. This group would attempt to clarify the components involved. These procedures will be discussed thoroughly under the chapter "Design of Treatments for Selected Skills."

C H A P T E R I I I

RESULTS OF THE SKILL IDENTIFICATION PROCEDURES

The following is the final list of skills needed for an individual to be a successful F-H evaluator. There are 164 skills on 58 levels of importance. It was not considered necessary to place each skill on its own level of importance, as the emphasis here was not on the creation of the perfect skills list but, rather, on the development of a fairly complete list in a rough priority order which was still specific enough so that the top few skills could be identified and used as a basis for the development and field testing of training materials. Those skills which came out of the Benedict documentation of F-H ("The Fortune/Hutchinson Evaluation Methodology, Version I, Draft I," September, 1973) are followed by "(Benedict, 1973a, __, __)" where the first blank refers to the major step in the methodology from which the skill came. The major steps will be represented by the following notation: I - Negotiation of Contract; II - Contract Decision Maker Reporting Process; III - Allocation of Resources; IV, I, II, IIA, IIB, III - The Goals Process, case I, II, IIA, IIB, III; V - Reporting Goals Data; VI, I, II, IIA, IIB, III - The Parts Process, case I, II, IIA, IIB, III; VII - Integration of Goals and Parts; VIII, I, II - Operationalization of Goals, case I, II; IX - Development of Observational Techniques; X - Implementation of Observational

Techniques; XI - Reporting Data; XII - Redesign; and XIII - Evaluation of the Evaluation. The second blank refers to the particular substep under the major step from which the skill was drawn, if that was the case. If, however, the skill was implied by a cluster of substeps or by the major step as a whole, the second blank is omitted. If the skill is found in more than one place in the Benedict documentation, only the first location is noted. Other notes following skills indicate skills drawn from related dissertation, e.g., "Jeffers, 1974, p. 31," "Gordon, 1972, p. 283." Items which are not referenced came from neither the Benedict documentation nor related dissertation.

List of Skills for F-H in Final Priority Order

1. The evaluator must be able to define evaluation (Benedict, 1973a, III, 1.21).
2. The evaluator should be able to deal with a lack of decision maker cooperation (Gordon, 1972, p. 279).
The evaluator needs to have and be able to use a variety of forms of the methodology including the shortest form, a short form, a long form, and the longest form for situations varying in resources available for evaluation (Rosen, 1974, p. 311).
The evaluator needs to have and be able to use a variety of forms for each part of the methodology (shortest, short, long, longest)--especially for operationalization of goals (Rosen, 1974, p. 311). (Perhaps a bank of goal operationalizations could be set up.)
3. The evaluator should be able to determine if there is a "real conflict" between the F-H definition of evaluation and the contract decision maker's definition (Benedict, 1973a, I, 1.6).
The evaluator should be able to determine whether a "mutual understanding is being maintained" (between evaluator and contract decision maker) (Benedict, 1973a, I, 3.1).

4. The evaluator should be able to proceed methodologically.
The evaluator should be able to communicate with the contract decision maker (Benedict, 1973a, I).
The evaluator should be able to deal with decision makers on a person-to-person level.
5. The evaluator should be able to utilize resources efficiently (Gordon, 1972, p. 282).
6. The evaluator should be able to determine when revisions in the definition of evaluation are necessary (Benedict, 1973a, I, 3.1).
The evaluator should be able to keep a group meeting from becoming non-positive and judgmental in nature.
The evaluator should be able to explain procedures to a group so that they will be understood (Benedict, 1973a, VI - IIB, 9.0).
7. The evaluator must be able to write coherent reports (Benedict, 1973a, II).
8. The evaluator should be able to furnish a broad outline of the methodology to the contract decision maker and others (Benedict, 1973a, I, 1.3).
The evaluator should be able to use the methodology as a framework and guide for evaluations.
The evaluator should have the ability to be humble when dealing with teachers to keep them from feeling threatened.
The evaluator should be able to deal effectively with individuals and groups.
The evaluator should have listening skills.
The evaluator should have communicating skills (with all decision makers).
The evaluator needs to be able to plan: what steps of what phases, parts, or sections of the methodology will be implemented for each decision maker; in what order; and with what minimum and maximum resources (Rosen, 1974, p. 314).
9. The evaluator should be able to explain methodological steps and logic to the contract decision maker and others (Benedict, 1973a, I).
The evaluator should understand and be able to explain the role of decision makers in evaluation (Benedict, 1973a, III, 1.2.2).

The evaluator should have person-to-person consultation skills.

10. The evaluator should be able to persevere in working on unpleasant tasks.
11. The evaluator must be able to communicate with others while doing tests of completeness (Benedict, 1973a, I, 4).

The evaluator must be able to pick out statements which appear to indicate what someone wants (the enterprise) to accomplish for oneself (or for others) (Benedict, 1973a, IV, 3.3).

The evaluator should be able to develop a hypothetical situation appropriate to the purpose of obtaining a decision maker's specific meaning for the goal in the context of the particular goal-part interface (Benedict, 1973a, VIII, 1.1).

The evaluator must be able to write down a stimulus in which the goal is absent (Benedict, 1973a, VIII - I, 1.5).

The evaluator needs skills in developing observational techniques (Gordon, 1972, p. 284).

The evaluator needs goal analysis skills (Benedict, 1973b, p. 206).

The evaluator needs skills to insure the goal dimensions are observed directly to their fullest extent (Gordon, 1972, p. 289).

The evaluator needs skills for deciding which observational techniques are most appropriate for which operationalized goals (Jeffers, 1974, p. 2).

The evaluator needs to be able to vary the directness, naturalness, and unobtrusiveness so that the observational technique can be maximally consistent with the decision maker's goals (Jeffers, 1974, p. 31).

The evaluator needs to be able to revise the evaluation if greater or fewer resources become available (Rosen, 1974, p. 314).

The evaluator should be able to schedule activities (Benedict, 1973a, XIII, 1.4).

12. The evaluator should fully understand the concepts of direct observation, natural conditions, and unobtrusiveness (Benedict, 1973a, IX, 4.0).

The evaluator should recognize the difficulties in the interpretation of the results of data collection (Benedict, 1973a, XI, 5.3).

The evaluator should know what to expect and what to look for with respect to deviations from the specified observational technique (Benedict, 1973a, X, 8.2).

The evaluator should understand the interaction of evaluation and decision making.

13. The evaluator should be able to "secure final approval and signature" of the contract decision maker or the "contracting group," if different from the holder of resources (Benedict, 1973a, I, 6.3).

14. The evaluator should have mastered certain assertive behaviors because decision makers must be led fairly forcibly at times.

The evaluator must know the procedures for amending the contract or letter of agreement (Benedict, 1973a, I).

The evaluator must know how to combine criteria to reach a final prioritization when more than one criterion is used to order a goals list (Benedict, 1973a, IV - IIA, 10.5).

The evaluator must be able to break down multiple responses to OFC (Benedict, 1973a, VIII - II, 6.0).

The evaluator needs to be able to insure that goal dimensions are fully operationalized.

15. The evaluator should understand how goals can be data for decision making (Benedict, 1973a, IV - III, 11.3).

The evaluator must know when a concept is "fuzzy," which concepts are "fuzzier" than others, and when a concept ceases to be "fuzzy" (Benedict, 1973a, VIII).

16. The evaluator needs to know how many resources are required for the evaluation to succeed (Benedict, 1973a, III, 3.2.3.1).

The evaluator should have an idea of how many resources in general are required to do sub-parts of major steps (Benedict, 1973a, IV).

17. The evaluator must know or be able to find out how the decision maker makes decisions (Benedict, 1973a, IV).

The evaluator should have skills in group relations.

The evaluator should be able to suggest several kinds of prioritization criteria for goals (Benedict, 1973a, IV - I, 9.0).

The evaluator must know and be able to explain the nature and purpose of the operationalization stimulus (Benedict, 1973a, VIII - I, 1.4).

The evaluator should be able to employ cost-effectiveness procedures for the proposed measurement plan and for the consideration of alternate plans (Jeffers, 1974, p.31).

18. The evaluator must know how to determine whether or not there are existing observational techniques that meet the plan (Benedict, 1973a, IX, 4.4).

The evaluator must know what "problems" are possible, i.e., what to look for in a field test of the observational technique (Benedict, 1973a, IX, 10.0).

The evaluator must know what constitutes "quality" data (Benedict, 1973a, X, 6.3).

The evaluator must be able to figure the degree of completeness of the operationalization of a goal (Benedict, 1973a, XI, 3.5).

The evaluator should recognize the consequences to the interpretation of data collected of the degree of operationalization performed (Benedict, 1973a, XI, 6.4).

19. The evaluator must understand how resources may be wasted by making careless decisions about the decision maker list.

The evaluator should have an idea of how many goals from how many different decision makers can be reasonably treated, given the resources for the evaluation (Benedict, 1973a, IV).

The evaluator needs to know possible costs of redesigning the evaluation (Benedict, 1973a, XII, 3.4).

20. The evaluator must know how to write down the positive ends of responses to the second hypothetical situation (Benedict, 1973a, VIII - II, 5.0).

The evaluator needs to be able to monitor the evaluation plan (Rosen, 1974, p. 314).

21. The evaluator must be able to identify threats to validity (of data collected or to be collected) (Benedict, 1973a, IX, 11.0).

The evaluator needs to know when and how sampling may threaten data quality (Benedict, 1973a, X, 6.3).

The evaluator should be able to document estimated loss of data quality due to the sampling plan (Benedict, 1973a, X, 6.6).

The evaluator should be able to document any deviations from the sampling plan (Benedict, 1973a, X, 8.3).

The evaluator should be able to document any other problems that occur (Benedict, 1973a, X, 8.4).

The evaluator needs to know when and how the number of observations threatens data quality (Benedict, 1973a, X, 6.4).

The evaluator needs to be able to deal with personal standards vs. objective judgment criteria (Jones, 1970, p. 199).

The evaluator should understand the constraints of contemporary schools through practical experience.

The evaluator should have a background in other evaluation approaches.

22. The evaluator should be able to see when the list of decision makers is "too large to be reasonable" in relation to the available resources (Benedict, 1973a, I, 5.3).

The evaluator must be able to re-allocate resources available, either among parts of the evaluation methodology or solely to the contract negotiation process (Benedict, 1973a, XII).

The evaluator needs to be able to make estimates of resources needed (for the evaluation, parts of it, etc.).

23. The evaluator should have skills in understanding group dynamics.

24. The evaluator should be able to eliminate overlap and redundancy in goal statements (Benedict, 1973a, IV - IIA, 8.3).

The evaluator must be able to note discrepancies in the activities test of completeness--e.g., activities with no relating goals and goals with no relating activities (Benedict, 1973a, IV - IIB, 17.5.1).

The evaluator should be able to prepare a discussion of goals-conclusions, implications, etc. (Benedict, 1973a, IV - III, 11.4).

The evaluator should have skills in logical thinking.

The evaluator should have sampling skills (Gordon, 1972, p. 283).

25. The evaluator must be able to know when a group size is small enough relative to the amount of resources available so that sampling is or is not necessary (Benedict, 1973a, IV - II, 2.0).

The evaluator should be able to determine comprehensiveness as a function of resources available (Gordon, 1972, p. 290).

The evaluator should be able to select valid, reliable observational techniques within the constraints of time and cost (Jeffers, 1974, p. 2).

The evaluator must know how many breakdowns of parts into subparts can be accomplished within available resources (Benedict, 1973a, V - IIA, 10.1).

The evaluator needs to know the possible benefits of redesigning the evaluation (Benedict, 1973a, XII, 3.3).

26. The evaluator must understand and be able to answer any decision maker questions about Fig. A, the Decision and Data Log (Rosen, 1974, p. 195).

The evaluator must be able to tell by looking at complete or partially completed DD log whether decision maker has misunderstood any part (Benedict, 1973a, III, 1.3.2.2).

27. The evaluator should be able to write general, global goals statements incorporating two or more goals (Benedict, 1973a, IV - IIB, 14.1).

The evaluator should be able to understand the fallibility of tests in general.

The evaluator should have skills in non-traditional measurement.

28. The evaluator should have public relations skills.

29. The evaluator should be able to see when the ability of the enterprise to deliver its objectives is jeopardized because too many resources are committed to the evaluation (Benedict, 1973a, I, 14.21).

The evaluator needs to be able to allocate resources to be used to design an alternative strategy satisfying the contract decision maker (Benedict, 1973a, II, 4.7).

The evaluator should know how limited resources affect the size of the sample and the sophistication of the sampling techniques (Benedict, 1973a, IV - IIB, 3.0).

The evaluator needs to know how limited resources will limit choice of prioritization criteria and the whole parts process (Benedict, 1973a, VI - IIB, 16.1).

30. The evaluator should understand and be able to use DRAC, DM schedule chart (DNSC), DSC and ORAC (Benedict, 1973a, VIII, 0.10).

The evaluator should be able to provide the decision maker with examples of conceptual components as major parts of different enterprises (without giving too many examples or examples of parts of the decision maker's enterprise and, therefore, leading the decision maker) (Benedict, 1973a, VI - I, 2.1).

The evaluator needs skills in adapting the methodology.

31. The evaluator should be able to present data in table format (Benedict, 1973a, XI, 3.9).

The evaluator should be able to present data in graph format (Benedict, 1973a, XI, 3.9).

The evaluator should be able to write behavioral objectives.

The evaluator should know how to write surveys.

The evaluator should have skills in questionnaire design.

The evaluator should have skills in training professionals and paraprofessionals to use observational techniques.

The evaluator should have conceptualizing skills.

32. The evaluator needs to be able to estimate how many extra resources would be required to develop novel reporting procedures (Benedict, 1973a, II, 3.2).

33. The evaluator should have skills from the classroom learning of the methodology.

34. The evaluator must be able to combine materials developed during the reporting period as a result of implementing the evaluation design (Benedict, 1973a, II, 5.2.4).

The evaluator should be able to apply problem-solving skills with individuals and groups.

The evaluator should have skills in clarifying processes.

35. The evaluator should be able to know when the matching of decision makers with resources is "realistic" (Benedict, 1973a, I, 5.62).

The evaluator needs to know how many resources are required to implement the standardized Monthly Reporting Process (Benedict, 1973a, II, 3.1).

The evaluator needs to be able to estimate the resources required to implement the End of Contract Period Reporting Process (Benedict, 1973a, II, 4.1).

36. The evaluator must be able to use Figures A, B, and C (Contract Decision Maker Reporting Process)(Benedict, 1973a, II).

The evaluator should be able to give examples which expand the contracting group's concept of the enterprise.

The evaluator should be able to give examples which limit the contracting group's concept of the enterprise.

The evaluator should be able to estimate a "short time" and a "long time," given the length of time in the evaluation contract (Benedict, 1973a, X, 10.4).

The evaluator should have a knowledge of sampling techniques and be able to draw samples (Benedict, 1973a, IV - IIB, 1.0).

The evaluator should know how to randomly assign goals lists to groups of decision makers and prioritization criteria to samples (Benedict, 1973a, IV - IIB, 10.1.5).

37. The evaluator should be able to choose a separate but similar enterprise from which to draw alternate goals lists (Benedict, 1973a, IV - IIB, 8.4.2).

The evaluator should have organizational skills.

The evaluator should be able to write test items.

The evaluator should have experience in filling methodological gaps.

38. The evaluator needs to be able to conduct a brainstorming session for identifying resources for the evaluation.

The evaluator should be able to employ cost-analysis procedures for the proposed sampling plan and consideration of alternate sampling plans (Jeffers, 1974, p. 32).

39. The evaluator should have skills from a previous implementation of F-H.

The evaluator must know what a frequency count is and how to compile one, e.g., for each goal on the list (Benedict, 1973a, IV - IIB, 10.2).

The evaluator should know how to generalize from samples, the dangers involved, etc. (Gordon, 1972, p. 283).

40. The evaluator should understand some content area well (outside of evaluation).

The evaluator should have analytical skills.

41. The evaluator must know how to compute a percentage of the number of members in the group who hold each goal on the list as a goal for the enterprise (Benedict, 1973a, IV - IIB, 10.2).

42. The evaluator needs to know how to identify decision makers who are likely to have parts other than the ones the decision makers the evaluator is working with are likely to put down (Benedict, 1973a, VI - IIB, 5.2).

43. The evaluator should have skills in making tables for reports.

44. The evaluator should be able to document the actual savings in resources due to the sampling plan.

45. The evaluator should have skills connected with "methodological thinking," i.e., going about a task with a singleness of purpose and a strong concern with order and precision.

The evaluator should have skills resulting from a more intuitive feel for the purpose and steps of the methodology.

46. The evaluator should be able to choose decision makers in the enterprise who are likely to have goals other than the ones the primary decision maker is likely to put down (Benedict, 1973a, IV, 4.2).

The evaluator must understand the characteristics of the group to be measured and be able to select a similar group (Benedict, 1973a, IX, 10.2).

The evaluator should be experienced in the role of being an administrator.

47. The evaluator should be able to determine when items on the resource lists are redundant or overlapping (Benedict, 1973a, I, 4.2).
48. The evaluator must be able to use the various charts, i.e., enter values in the right places, perform correct computations, etc. (Benedict, 1973a, XIII, 2.6).
The evaluator should have skills in developing attitude surveys.
49. The evaluator should have skills from having done methodological development.
50. The evaluator should be able to tolerate ambiguity.
The evaluator should have skills in item sampling.
51. The evaluator should be familiar with and able to use Meta-Methodology (Benedict, 1973a, II, 3.8.3).
52. The evaluator needs to be able to determine the reliability and validity for the proposed measurement for the sample (Jeffers, 1974, p. 32).
53. The evaluator should know and be able to use the techniques of inquiry.
The evaluator should have skills in curriculum development.
54. The evaluator should be able to develop a complete plan for sampling from the population of observations (Benedict, 1973a, X, 6.5).
The evaluator should be able to use a formula such as:
$$E = \frac{NDU}{NDR} \times 100$$
 (Benedict, 1973a, XIII, 2.6).
The evaluator should be able to add.
The evaluator must be able to compute percentages of total numbers of hours (and other resources) (Benedict, 1973a, III, 4).
The evaluator must know how to multiply (Benedict, 1973a, IV - III, 10.4).

The evaluator should know how to count (Benedict, 1973a, IV - IIB, 11.0).

55. The evaluator should know statistics.

56. The evaluation should have skills in business administration.

The evaluator should understand the basic principles relating to the economics of education.

57. The evaluator should be able to statistically adjust a set of observations made using the actual measurement technique to more fully coincide with those made using the ideal measurement technique, if possible (Benedict, 1973a, IX, 10.4.3).

58. The evaluator should know computer programming.

The evaluator should be able to use computer programs for evaluation purposes.

Discussion of the Skills List

For the final skills list, the most important source of skills was found to be the Benedict documentation of F-H which furnished 52% of the skills (all percentages are rounded to the nearest whole percents). The questionnaires given to graduate students accounted for 27%, related dissertations for 15%, and other sources such as articles on evaluation, other models, etc., for only 6%. It was expected that the most complete documentation available would be the source of most of the skills needed for F-H since all procedures used in F-H are included in the documentation. This source did not, however, supply the same percentage of skills in all six categories described in the previous chapter. It supplied 57% of skills under "Knowledge of the Methodology,"

52% of skills under "Other Knowledge," 68% of "Resource Skills," and 65% of "Statistical and Quantitative Skills," but only 37% of "Interpersonal Skills" and 20% of skills under "Knowledge from Experience." The questionnaires, on the other hand, supplied 58% of the "Interpersonal Skills," 53% of "Knowledge from Experience," and 33% of "Other Knowledge," while only 4% of "Resource Skills," came from this source. Related dissertations furnished 36% of skills under "Knowledge of the Methodology," and minor percentages in the other categories. Other sources provided none of the skills under "Knowledge of the Methodology," and "Interpersonal Skills," and minor percentages in the other categories.

It was expected that skills under "Knowledge of the Methodology" would not come from other sources, especially because of the completeness of the Benedict documentation. The fact that these sources did not provide more skills in other categories could have been because it is not really known what evaluators do or the skills they should have, as was suggested by Guba and Gephart (1970). A contributing reason could also be that the least amount of time was spent on the search of other sources because it was not expected to be fruitful. This may have been in part self-fulfilling.

Table I is interesting because the responses of the graduate students to the questionnaire indicate their emphasis on skills which are needed for using F-H but which are not found in the documentation. These were skills which would be very useful when dealing with educational decision

TABLE I
NUMBER AND PERCENTAGE OF SKILLS ON FINAL LIST FROM
EACH SOURCE BY CATEGORY AND TOTAL

Skill Source	Number and % of Skills Under:						Number and % of Total Skills from this Source
	"Knowledge of the Methodol- ogy"	"Knowledge from Experi- ence"	"Resource Skills"	"Interper- sonal Skills"	"Statisti- cal and Quantita- tive"	"Other Knowl- edge"	
Benedict	161	3	19	7	11	29	85
Documentation	192	4	22	8	13	34	100
	573	20	68	37	65	52	-
	104	2	12	4	7	18	52
Questionnaire	2	8	1	11	3	19	44
	5	18	2	25	7	43	100
	7	53	4	58	18	33	-
	1	5	0	7	2	12	27
Related	10	3	5	1	2	3	24
Dissertations	42	12	21	4	8	13	100
	36	20	18	5	12	5	-
	16	2	3	0	1	2	15
Other Sources	0	1	3	0	1	6	11
	0	9	27	0	9	55	100
	0	7	10	0	5	10	-
	0	0	2	0	0	4	6
Total Count	28	15	28	19	17	57	164
and Percent	-	-	-	-	-	-	100
of Total	100	100	100	100	100	100	100
	17	9	17	12	10	35	100

¹Count; ²Row Percent; ³Column Percent; ⁴Percent of the Total.

makers, e.g., skills in group dynamics, problem solving techniques, and communications. It was expected that the bulk of skills required by an F-H evaluator would be specified in the documentation itself and in related dissertations, and, in fact, these two sources supplied 93% of the skills in the category "Knowledge of the Methodology." The responses to the questionnaire also indicate the absence of concern about statistical and resource skills on the part of the students. Apparently they felt that these skills were not too important for using F-H successfully.

This conclusion was further supported by the order of skills on the final list. When the skills list was divided roughly into thirds on the basis of importance (priority levels #1-18, or 37% of the skills; priority levels #19-34, or 32% of the skills; and priority levels #35-58, or 31% of the skills), it was found that all of the "Statistical and Quantitative Skills" fell into the bottom "third." The reader will recall that the final list resulted from combining and weighting the three skills lists ordered on the criterion of "importance for being a successful F-H evaluator" by this author, Dr. Hutchinson, and Dr. Benedict. Thus, these three persons also tended to discount the importance of statistical skills to F-H.

The most important categories turned out to be "Interpersonal Skills," with 89% of these falling into the top "third" of the final list, and "Knowledge of the Methodology," with 81% of these falling into the top "third." Thus, the combined list indicates that a successful F-H evaluator

needs to be particularly adept at dealing with people and to have mastered the methodology itself, in the view of this author, Dr. Hutchinson, and Dr. Benedict.

Of the remaining three categories, 57% of the skills under "Knowledge from Experience" fell into the bottom "third." This apparently indicates that these skills are less important than those in the two categories just discussed. Skills under "Resources" and "Other Knowledge" fell fairly heavily into the middle "third" with 57% of the former and 46% of the latter on this level. Remaining skills under both categories were split fairly evenly into the top and bottom "thirds."

Finally, it should be noted exactly what percentage of the final list each category constituted. "Other Knowledge" was the largest with 35%, or 57 skills. "Knowledge of the Methodology" and "Resource Skills" each made up 17% with 28 skills each. The smallest groups were "Interpersonal Skills" with 12%, or 19 skills, "Statistical and Quantitative Skills" with 10%, or 17 skills, and "Knowledge from Experience" with 9%, or 15 skills.

Methodological Inadequacies

There were a few problems with the way that the skills list was developed which could be corrected, in order to improve the final product. The way the skills were extracted from the documentation of the methodology, related dissertations, and other sources seems sound. The questions asked to the graduate students, however, might have been improved.

The second question, "What skills do you feel that you developed by learning the methodology"? was interpreted by some as meaning those skills one obtained by the classroom study of the methodology. Others answered the question with respect to those skills gained by the actual application of the methodology. A few seemed to respond to both of these interpretations. This process could be improved by adding the additional question "What skills did you develop by the application of the methodology?" This would insure that the skills list elicited by this means would be as complete as possible. In addition, it might be a good idea to emphasize that the responses to all questions should include relatively unimportant, as well as the important, skills.

Secondly, there were some inaccuracies in the final list, because of the categories used. A skill such as "the evaluator should be able to select valid, reliable observational techniques within the constraints of time and cost" could have fitted into two categories--"Knowledge of the Methodology" and "Resource Skills" It was placed into the second one because this author felt the emphasis was on the resource aspect of the skill. An argument could be made for the first category, however. Because the final list came about through a combination of priority levels of both the individual skills and their respective categories, if a skill was placed in a category incorrectly, it would affect its placement on the final list. Thus, the categories could be

made more specific so that this problem would be avoided and the final list could be in a more accurate priority order.

Implications for Further Research

The development of a very complete skills list has substantive implications for curriculum planning for F-H. Now the curriculum developer has a choice about just what will be taught. (S)he can plan a course based on a "nested" approach to the methodology, i.e., covering the most important objectives for each major step of F-H first and then becoming more and more specific, depending upon how much skill the students wish to have. The developer can also concentrate on the various categories of skills, and, as the responses to the questionnaire seemed to indicate, much more needs to be done in the area of interpersonal skills. Finally, this skills list may help those who plan courses of study for future professional evaluators. The statistical and quantitative skills could be taught in a course separate from the F-H course, possibly as a pre-requisite or a co-requirement, and those skills under "Other Knowledge" could be handled under a variety of courses.

C H A P T E R I V

DESIGN OF TREATMENTS AND TESTS FOR SELECTED SKILLS

Skill Selection

Two skills having the highest priorities on the final list were, "The evaluator must be able to define evaluation," the only skill on the top priority level, and "The evaluator should be able to deal with a lack of decision maker cooperation," which was one of three skills on the second priority level. It was decided that curriculum materials for training evaluators in these skills would be developed and field tested. The skill concerning decision maker cooperation was chosen over the other two skills on the second priority level. The latter concerned the evaluator's ability to use a variety of forms of the methodology, given different resources and with different parts of the methodology. These two skills would seem to be better handled through methodological development. In this way, the various forms of F-H would be built into the documented methodology itself, and the future evaluator could be directed to use a particular form of the methodology, rather than to be taught to use all forms by means of evaluator training procedures. This would seem to make better use of the teacher's and student's time. Because approximately half of the author's time to be devoted to the activities required by this dissertation had been expended during the skill identification process, curriculum and field tests could be designed and carried out for no more than two skills.

General Approach to the Design of Treatments and Tests

The approach to the development of instructional materials used here followed the guidelines put forth by Tyler (1950) and discussed in Chapter I. For each skill, instructional objectives had to be specified. This would be done by a careful consideration of the behaviors which are evident in a successful F-H evaluator who is competent in the skill. From there, learning experiences were structured to teach the objectives, and appropriate test items designed. Responses to the latter, after instruction, would determine the student's competency or lack of competency for the skill. The learning experiences were also designed in such a way that the student would be able to practice the behaviors elicited by the objectives and be reinforced in some way for achieving them (Skinner, 1968). The students' anticipated entering behaviors also affected the design of the learning experiences, since their time would be wasted learning objectives they already knew or attempting objectives for which they did not have the necessary prerequisite knowledge and/or skills.

Selected Mode of Instruction

The mode of instruction selected for both skills was the self-instructional booklet. This choice seemed attractive because this author's present skills could be used, and it was not necessary that she learn a whole new field from the beginning, such as film-making. Determining the student's achievement or lack of achievement of competency in the skill

was not a major problem with this format as with some of the others. Also, the instructional process was replicable because of the standardized format. In addition, the use of review questions throughout the workbooks could allow the students to practice the behaviors being taught and be immediately reinforced for correct answers.

Other possible devices had included lecture, small group investigation, simulation, slide-tape presentation, movie, and overhead transparencies. The slide-tape and movie presentations were eliminated because of time constraints and this author's desire to focus the dissertation on the content area of educational evaluation, rather than on the development of an intriguing mode of instruction. The lecture-format, in this author's opinion, has traditionally been a poor way of teaching a significant amount of new material, as it does not allow for student participation in the learning process, as is advocated by Tyler (1950), Skinner (1968), and others. Overhead transparencies were seen by this author as merely a means of embellishing the lecture format, rather than as representing a significantly different mode of instruction. The small group investigation might have been very effective, but the training of a facilitator for each group and the difficulty of evaluating individual student competency in this setting would have been difficult problems to overcome. Finally, the difficulty of evaluation would also have been a problem with the simulation or role-playing

approach, although a taped role playing situation might also be very effective as a teaching device, as well as a test of student competency in the skill.

Design of Treatment and Test for "Defining Evaluation"

The booklet covering the skill "The evaluator should be able to define evaluation" was directed toward the beginning student of F-H. The author tried to make it broad-based so that it might have other uses, e.g., with classroom teachers who would like to better evaluate their instruction or with program decision makers who might be potential clients for an F-H evaluation.

Objectives. The behaviors to be brought about by attaining competency in this skill, or the instructional objectives, were defined as follows:

1. The student should be able to give the F-H definition of evaluation.
2. The student should be able to recognize the assumptions behind the F-H definition of evaluation.
3. The student should be able to give the most important implication of the F-H definition of evaluation.
4. The student should be able to name the three criteria to judge the effectiveness of an F-H evaluation.
5. The student should be able to define the three criteria.
6. The student should be able to recognize the

processes in F-H which attempt to insure that these three criteria are met.

7. The student should be able to recognize the major problems behind some other evaluation models.
8. The student should be able to give the major disadvantage of models, as compared to methodologies.
9. The student should be able to answer decision maker questions by explaining: the roles of evaluator and decision maker in an F-H evaluation; that F-H is not the same thing as public relations, needs analysis, decision making, research, or measurement; and that some models do have some similarities to F-H.

It can be seen, then, that an ostensibly simple skill such as "defining evaluation" implies objectives which go far beyond the first one which is to be able to give the F-H definition of evaluation. The question of whether or not so many objectives should have been attempted in one booklet, as well as other issues pertaining to the appropriateness of the content, correspondence between content and test items, etc., will be discussed in the next chapter.

General structure of the booklet. The booklet in the form which was field tested can be found in Appendix A, along with the final review, and the facilitator's manual. The booklet itself is thirty-one pages long and contains an introduction and instructions to the reader in addition to

the content material which is interspersed with review questions. There were never more than two pages of content material before each review question, and there were a total of nine review questions in all which the reader was asked to answer right in the booklet. It was assumed that the reader would require about one hour to go through the booklet and answer the questions.

General content of the booklet. The content of the booklet was designed to provide an interested new-comer to the field of evaluation with some background in one form of evaluation, F-H, and an idea about what F-H can and can't do. The booklet includes in the following order: the F-H definition of evaluation; the assumptions behind the definition; three other reasons frequently given for doing an evaluation; very brief descriptions of four other evaluation models; a brief discussion of why an F-H evaluation should not be equated with either research or simply measurement; the most important implication of the purpose of F-H, the resulting three criteria for judging an F-H evaluation, and their definitions; and the three kinds of processes which are built into F-H to help insure that the three criteria are met. These content areas relate directly to the list of objectives for this skill given above. The order given was chosen because it seemed to this author to represent a logical sequence for the introduction to the topic, answering the questions "What is F-H?", "How is it different from other

means of evaluation?", "What are the implications of using it?", and "How does F-H attempt to fulfill its purpose?"

Review questions. Of the nine review questions in the booklet, two were multiple choice. These questions required the reader to pick the F-H definition of evaluation out of four choices, and to determine what it was that the goals and parts processes, tests of completeness, and putting lists in priority order helped to insure, again out of four choices. Four questions required matching. They were: matching four evaluation models with their appropriate characteristics; taking a list of nine statements such as "random assignment" and "many interacting variables" and matching them with either "research" or "evaluation"; matching the three criteria for a successful F-H evaluation with an appropriate description; and matching the same three criteria with the process in F-H designed to accomplish each one. The other three questions asked for a response from the reader without any assistance. These included: answering the question about why F-H should not be equated with public relations, needs analysis, and decision making; giving four assumptions behind the F-H definition of evaluation; and answering the question about what is the specific source of observational techniques in F-H.

It was hoped that the variety of question formats would be stimulating to the reader, although this might also complicate the interpretation of test scores. Because extensive field testing of all materials was planned before classroom applications would be attempted, however, the effect of this

variety of formats would be thoroughly examined in advance. In each case, the review question immediately followed the pertinent content material in the booklet. This format has an advantage over others in that an individual can be reinforced immediately upon answering the question by turning the page and discovering whether or not his or her answer was correct. While answering correctly is probably reinforcement enough, the booklet also offers short words of commendation such as "Good work," or "Well done" for correct answers.

Final review. The final review consisted of eight questions, some with several parts. The first question asked the reader to give the F-H definition of evaluation, without assistance. This is more difficult than the corresponding review question which asked the reader to choose the definition out of four choices. The second question asked the reader to select the assumptions behind the F-H definition of evaluation out of a list of fourteen possible assumptions. This was probably easier than the corresponding review question which asked the reader to write down four of the assumptions behind F-H without any assistance. The next question gave the reader seven hypothetical statements from decision makers and asked him or her to respond in the role of an F-H evaluator. This caused the reader to draw on the content from the booklet as a whole and tried to get at the reader's overall understanding, rather than at some memorized term. Question four had two parts, each of which asked the reader to select from four choices the one major problem with the

evaluation model indicated. The next question asked the reader, without assistance, to name the major problem with models, as opposed to methodologies, in general, and question six asked the reader to give a similarity between F-H and two other evaluation models, again without assistance. These questions called for memory of some content material. The last two questions were concerned with the three criteria for a successful evaluation. Question seven asked the reader to list the three criteria and define them, again from memory alone. The last question gave the names of three processes associated with the criteria. After each one were four statements and the reader was asked to check off all those that applied. For each process, the correct choices were the criterion that each helped the evaluation to meet and a one-line description of the process itself. The eight questions of the final review, then, attempted to test the objectives listed above which assess different levels of learning and understanding. The reader was given about half an hour to complete the test and was asked not to refer back to the booklet. The reader was not able to correct the final review alone, and, therefore, reinforcement was delayed until the facilitator could correct it and return it to the reader.

Facilitator's handbook. This handbook was designed for someone who is familiar with F-H to use to determine whether or not those individuals who had taken the final review should be judged competent in the skill "Defining evaluation." The handbook included general instructions to the facilitator on the administration of the booklet and the

final review, permitting him or her to answer questions while students were reading the booklet and asking that he or she see that booklets are closed during the final review. The correct answers for the matching, multiple-choice, and fill-in questions were included in the booklet, along with sample responses to the open-ended questions asking students to answer hypothetical decision maker questions. Points were assigned to each question with a certain minimum number of points required for questions one, two, three, six, seven, and eight in order for competency to be attained. Questions four and five which dealt with other evaluation models and models in general were not considered necessary for competency. The total number of possible points was 42. If a student achieved less than 29, he or she was asked to review the entire booklet and to take the final review again. If a student failed to earn the necessary number of points on certain questions but gained more than 29 points over-all, the facilitator was to ask him or her to review selected pages in the booklet, as they are listed in the facilitator's handbook under "Remedial Materials."

Design of Treatment and Test for "Dealing with a Lack of Decision Maker Cooperation"

Defining the skill. Because the skill "Dealing with a lack of decision maker cooperation" is more complex than the previous skill discussed above, some time had to be spent

defining the parameters to be covered before the objectives could be stated. It was decided that "dealing with" should be operationalized as recognizing symptoms, diagnosing the most likely problem, and choosing an appropriate course of action. In order to determine the problems that would be covered in this workbook, this author met with two other graduate students who were familiar with and had used F-H. This group brainstormed the sets of symptoms, problems, and courses of action with which they had become acquainted through their own experiences. These three lists were next presented to three other graduate students who were also very experienced in the use of F-H so that they could react to them and add anything else that they happened to think would be important to include. This second group thereby acted as a "test of completeness" to the first group, i.e., the second group helped to insure that the original lists became as complete as possible.

The second list, that of possible problems, was examined for vagueness and overlap. After collapsing the list and eliminating vagueness, ten possible problems emerged. It was decided to treat them as discrete entities for the purpose of instruction, although in the real evaluation situation it is likely that an uncooperative decision maker may embody more than one problem. Finally, it was decided that the third list, "Courses of action," should be divided into two lists--ways of avoiding the problem before it occurs and ways of

alleviating the problem once it has developed. At this point objectives could be stated.

Objectives. This booklet was designed to be used by future F-H evaluators who have learned the major steps in the methodology but who have not yet tried to apply it. The instructional objectives are as follows:

- A The student should be able to recognize the symptoms for the following problems associated with a lack of decision maker cooperation, and, for each problem, be able to give one or more ways of avoiding the problem and one or more ways of alleviating it.
- 1) The decision maker has a lack of time.
 - 2) The decision maker doesn't want to know what he or she is doing wrong.
 - 3) The decision maker lacks an understanding of the methodology.
 - 4) The decision maker has philosophical disagreements.
 - 5) The decision maker is not "methodologically oriented."
 - 6) The decision maker has a "status hang-up."
 - 7) The decision maker is impatient with the initial processes.
 - 8) The decision maker is insecure or under pressure.

- 9) The documentation may be dull, dry, or threatening.
- 10) The decision maker was wrongly identified by the contract decision maker.
- B The student should be able to give a general approach for dealing with a lack of decision maker cooperation.
- C The student should understand that the most likely problem may not be the one really causing the lack of decision maker cooperation. For each problem listed above, the student should be able to give another possible problem causing the troublesome symptoms.

General structure of the booklet. The booklet in the form which was field tested can be found in Appendix B, along with the final review and the facilitator's manual. The booklet is sixty-two pages long and contains an introduction and instructions to the reader preceding the content material. The material is interspersed with review questions, and there is a maximum of four pages of content between review questions. Following the last review, there are two summaries. The first summarizes all the precautions an evaluator might take to avoid the problems covered in the booklet, and the second summarizes the general approach an evaluator might use to deal with a lack of decision maker cooperation. There were a total of eleven sets of review questions, and it was

assumed that the reader would require at least one hour to go through the booklet.

General content of the booklet. Contained in the booklet are ten possible ways a decision maker might reveal a lack of cooperation to an evaluator. These are in scenario form with the reader ("you") included as the evaluator. Each scenario is followed by a probable diagnosis, a list of ways to prevent the problem in the future, a list of ways to alleviate the problem, and a list of other problems to investigate in case the one given in the diagnosis, the apparent problem, is not the right one. The following is the list of problems in the order in which they appear in the booklet: the decision maker lacks time; the decision maker doesn't want to know what he or she is doing wrong; the decision maker lacks an understanding of the methodology; the decision maker has philosophical disagreements; the decision maker is not "methodologically oriented"; the decision maker has a "status hang-up"; the decision maker is impatient with the initial processes; the decision maker is insecure and/or under pressure; the documentation of the methodology may be dull, dry, or threatening; and the decision maker was wrongfully identified by the contract decision maker. This order represents this author's interpretation of the importance of the problems, i.e., their likelihood of occurrence. It can be seen that this content material and the two summaries discussed above are directly related to the instructional objectives.

Review questions. In this booklet, the eleven sets of review questions were almost uniform in format, unlike those for "Defining evaluation." The first question asked about the general procedures for dealing with a lack of decision maker cooperation which were very briefly summarized in the introduction. The reader was asked to give one thing an evaluator could do when efforts to alleviate a lack of cooperation had failed. All of the other review problem sets asked for the reader to write down, without assistance, one or two ways to avoid the problem just discussed and/or one or two ways to alleviate the problem. In addition, three of the problem sets briefly summarized the symptoms of the problem just covered and asked the reader to fill in the problem name. Finally, in the tenth problem set, the first question asked the reader to underline the phrase ("difficult," "somewhat difficult," "fairly easy") which best described the process of avoiding the problem of documentation which may be dull, dry, or threatening. Again, the review questions immediately followed the pertinent content material, and, again, the reader could be immediately reinforced for answering correctly by turning the page of the booklet and being able to read the correct answer which was usually followed by some words of commendation such as "Well done" or "Good job."

Final review. The final review consisted of six questions, two of which had more than one part. The first question had three parts. In it three scenarios were given, and

for each one the reader was asked to recognize the most likely problem embodied in the scenario; to give one way of avoiding this problem; to give one way of alleviating it; and to give another possible problem which might be causing these decision maker behaviors. The second question involved matching a set of symptoms with the appropriate problem. Nine sets of symptoms were given for only seven problems, so that two sets of symptoms had no match. These two questions related to the set of objectives under "A" and, in part, to the objective under "C." The third question asked the reader to name the most common problem, and the fourth asked for the name of the problem which is the most difficult to diagnose but which often underlies several other problems. Knowing the answers to these will help the student achieve the instructional objectives under "A" and "C." The fifth question asked the reader to write down seven things an evaluator can do in the beginning to avoid problems. The reader should remember these from the summary given at the end of the booklet. Again, this question related to the objectives under "A" above. The last question asked the reader to write down four general steps that an evaluator might use to deal with a lack of decision maker cooperation. These steps were also included in a summary at the end of the booklet and evaluate the achievement of the instructional objectives under "B" above.

It was estimated that this test would take about half an hour to complete. Again, the students were asked not to

refer back to the booklets, and the tests were corrected by the facilitator so that reinforcement of correct answers was delayed.

Facilitator's handbook. The facilitator's handbook was designed to allow someone familiar with F-H to determine whether or not students were competent in "Dealing with a lack of decision maker cooperation" after having taken the final review. The handbook included general instructions to the facilitator and the specific answers desired for the questions on the final review. For questions five and six, the facilitator was referred to the summaries at the end of the booklet itself.

There was a minimum number of points required on each question in order for a student to achieve competency in this skill. The maximum number of points overall was 34. If a student earned less than 25 points, he or she should review the entire booklet and take the test again. If a student failed to achieve the minimum on one or more questions, however, but gained more than 25 points in all, the facilitator should refer him or her to the appropriate pages listed in the handbook under "Remedial work."

C H A P T E R V

FIELD TESTS OF MATERIALS FOR SELECTED SKILLS

Design of the Field Test for "Defining Evaluation"

The instructional workbook designed to teach the skill "Defining evaluation" was intended for newcomers to the field who are interested in learning about F-H. It was hoped that even those unfamiliar with the concepts and terminology of evaluation would be able to understand the content of this workbook, thereby gaining an introduction to F-H. It was also hoped that this introduction would encourage them to pursue the topic further, although the achievement of this goal was not going to be measured during this field test because of this author's desire to focus the available resources on the previous objective.

Because the purpose of the first field test of instructional materials is to identify major problems which will be subsequently revised, it is important to choose a small number of individuals who would best be able to articulate any difficulties that they have with the booklet. While a large number of participants may be able to identify more problems than a small number, a parsimonious use of time dictates that the first field test should focus on major problems only which should be identified as easily by a small group as by a large one (Thomann, 1973).

Two groups were therefore selected as likely candidates for the field test of this workbook. The first consisted of those people who worked in the area of evaluation for the Special programs at the University of Massachusetts in Amherst. This author had worked with them during the past year and a half and had introduced them to a very generalized version of F-H without having mentioned it by name or having taught it in its entirety. This group included full-time graduate students who worked twenty hours a week for one program or another, as well as the directors of the programs who worked full-time in this capacity. None in this group had had the Evaluation Methodology I course, although four had had some experience in educational evaluation, especially with goal identification and questionnaire construction. Both graduate students and directors were generally responsible for many more tasks than just evaluation activities, and overall they were beginners in the field.

The second group consisted of students taking the course "Evaluation Practicum" taught by Dr. Hutchinson at the University of Massachusetts at Amherst. This course combined instruction in F-H with actual field experiences in evaluation. Because it was the beginning of the course, however, these students had not yet learned much about F-H and were considered to be beginners.

Members of the Special Programs group were contacted by this author by telephone and six agreed to participate. The date decided upon was February 25, 1975, from 12:30 to 2:00 p.m. It was estimated that it would take them an hour to

read through the workbook and about half an hour for the final review. The field test took place in the informal atmosphere of a classroom in one of the dormitories on campus. There was no discussion during the reading of the booklet and the completion of the final review, but many valuable comments were brought out and recorded by the author at the end of the session.

Three of the five members of Dr. Hutchinson's class who were present on February 20, 1975, went through the workbook and took the final review at this time. The other two were given the workbook on February 24, 1975, to do at home, and they took the final review in class on March 3, 1975.

Background questions. No pre-test was designed for this skill because it was decided that it would serve no purpose for this kind of field test. This field test was not done to monitor the individual's increase in knowledge, but to identify problems with the workbook. Since only two participants became competent in this skill, several problems were, indeed, identified through an analysis of the results. This also showed that the groups selected for the field test on the assumption that they would not be competent were correctly identified, proving ex post facto that a pre-test was not required.

A short collection of background questions was put together for the purpose of determining whether the differences in the scores on the final review might be attributable to differences in background, training, and/or experience in

the field of educational evaluation. These questions and those designed to elicit the participants' reactions to the materials discussed below under "Student reaction" can be found in Appendix C. The questions included: whether or not the individual had taken the "Evaluation Methodology I" course at the School of Education, University of Massachusetts at Amherst (this course teaches the specific steps of F-H and is often taken before "Evaluation Practicum"); the present student status of the individual; whether or not the individual had had experience doing educational evaluation; and whether or not the individual was "familiar with" F-H.

Student reaction. It was also desirable to find out what those people using the booklet thought about the content, the review questions, and the final review. The reader was asked to consider the material in the booklet and to respond on a five-point scale to the following three variables: from "easy to follow" to "difficult to understand"; from "clearly stated" to "confusing or jargony"; and from "was comfortably completed within the given time" to "took too long." The reader was also asked to consider the set of review questions and the final review, and, for each set of questions, to circle all that applied of the following: "too easy"; "too difficult"; "challenging"; "too time-consuming"; "appropriate to the workbook content"; and "other (please specify)." The last question asked the reader to make any other comments about the content, the review questions,

and/or the final review that (s)he wished. These results were to be used eventually to revise and improve the booklet.

Experts. In addition to the eleven participants in the two groups described above, it was decided that the final review, without the workbook, should be given to students of F-H who had been trained in the methodology and had also used it in a practical setting. This was done to verify whether or not these "experts" had become competent in this skill through previous training and experience, even though they had not been trained in "Defining evaluation" in this specific way. It was unknown whether this group would be competent in this skill, but the real purpose of using them was to make sure that the workbook was a necessary addition to existing training procedures. Two graduate students who worked in the Student Affairs Research and Evaluation Office at the University of Massachusetts at Amherst and the director of that office agreed to take the test which they did individually during the week of March 10, 1975. In addition, another graduate student working for the "Clinic to Improve University Teaching," also at the University of Massachusetts, took this test during the week of March 17, 1975.

Results and Discussion of the Field Test of "Defining Evaluation"

The completion of the workbook took about half an hour which was about half the anticipated time. The final review

was completed in about twenty to thirty-five minutes. No questions were asked during the reading of the workbook or the completion of the final review.

Review questions. For the most part, there were few problems with the nine review questions distributed throughout the workbook. No one missed more than three questions, and nine out of the eleven participants missed two or fewer. In general, if a participant misses a review question and gets the corresponding question on the final review correct, it is assumed that some learning has taken place, providing that the questions really are sampling the same content. If a reader answers the review question correctly and misses the corresponding final review question, the requirements for competency in the skill should be re-examined which may lead to changes in the content and/or the final review question.

Four of the questions were answered correctly by all participants, and three more were missed by only one out of eleven. Question number two, however, did cause problems for six out of eleven participants. This question asked the reader to list four assumptions behind the F-H definition of evaluation after having read through a list of seven assumptions on previous pages. These assumptions had not been explained at all in the workbook. One reader had omitted the question, one missed three out of four assumptions, two missed two assumptions, and two missed only one. Five out of eleven answered the question correctly

An issue which should be considered is whether or not the evaluator in the field must be able to give these specific assumptions from memory or only be able to give a general rationale for F-H. In the booklet, the review question was followed on the next page by the list of seven assumptions which provided the reader with a second opportunity to review them. This may have been sufficient for the student to learn these assumptions, as will be discussed in connection with the second question of the final review which asked the reader to select the assumptions from a list of fourteen possible ones.

Another question which could promote a possible reexamination of the competency requirements for this skill was review question number five. In this question, the reader was given a list of nine descriptive phrases and asked to identify the ones which applied to classical research design and those applying to evaluation, again after having read only a brief paragraph contrasting the two fields. Five out of eleven participants had some kind of problem with this question. One omitted it and one got it completely wrong, i.e., matched all the phrases describing classical research design with evaluation, and vice versa. Of the other three, all matched "treatment changes" with classical research design rather than evaluation, and two out of the three matched "holding treatment intact" with evaluation, instead of classical research design. Although these results do not

indicate a definite problem with this question, it should be considered whether or not the question is appropriate to the audience for whom the workbook is intended. In addition, the purpose of the workbook is to teach the reader about defining evaluation. Thus, the review question would be more consistent with this if it asked the reader merely to pick out the terms which belong with evaluation leaving classical research design out of it. A chart showing these connections should be added to supplement the present workbook content which could be made somewhat more explanatory itself. The related question on the final review could ask the reader to list three or four general attributes of evaluation, as opposed to research. The present final review question only listed "An evaluation should make as much use as possible of classical research design" as one of a list of possible assumptions behind F-H to which the reader was asked to respond. Thus, in its present form, the final review question is much simpler than the related review question and only asks that the reader have learned that the above is not one of the assumptions behind the F-H definition of evaluation.

Final review. Of the eleven persons in the field test group who took the final review, only two were judged competent in this skill. This discussion will, therefore, focus on an investigation of the achievement of competency for each of the eight individual questions in the final review. Table

TABLE II

"DEFINING EVALUATION": FINAL REVIEW QUESTION COMPETENCY ACHIEVED BY
THE FIELD TEST GROUP AND THE "EXPERTS"

Final Review Question #	Field Test Group											"Experts"			
	Special Programs						Hutchinson's Class								
	#1	#2	#3*	#4	#5	#6	#7*	#8	#9	#10	#11	#1	#2	#3	#4
1			X	X		X	X		X			X	X	X	X
2	X	X	X	X	X	X	X	X	X	X		X		X	X
3			X		X	X	X					X	X		
6	X	X	X	X	X	X	X					X	X	X	X
7		X	X			X	X						X		X
8	X		X		X		X					X	X	X	

*Competent in the skill.

Note: Questions four and five had no minimum number of points required for competency.

II indicates the scores of the field test group and the "experts" on the final review and the pertinent variables discussed above under "Background questions." An examination of these variables will occur later in this chapter.

In general, the results for the participants from Special Programs were similar to those for Dr. Hutchinson's class, as can be seen in Table II. Therefore, and because of the small numbers in each group, these groups will be considered together. The "experts" will be discussed separately in a later section of this chapter in order to determine how necessary the workbook is for evaluator training, as well as to investigate any unusual differences in the response patterns of the "experts" vis à vis the field test group.

The first question asked the reader to give the F-H definition of evaluation. Six field test participants (55%) answered correctly. Of the five who missed it, one had completely omitted it. The other four all had the idea that an evaluation provides data for decision makers, but they neglected to say that the data was actually to be used for their decision making. This distinction was not stressed in the workbook, yet it is an important one. It is the use of the data by decision makers which is the strongest concern of F-H evaluators. This result implies that the material in the workbook needs to be expanded in this regard.

The second question asked that the reader check off on a list of fourteen possible assumptions all those that really

were assumptions behind the F-H definition of evaluation. There were seven assumptions which should have been identified, but the reader was not given this information. All field test participants were competent on this question.

It would seem, then, that these results indicate that the instructional objective relating to this question was being met very well. The only problem is that the corresponding review question in the workbook which asked the reader to list four assumptions behind F-H from memory might have caused unnecessary anxiety and resistance. Some people resent being asked to parrot newly-learned information. Because an evaluator should at some time, however, have at least some of these assumptions at his/her fingertips, it would probably be more appropriate to make this question the one in the workbook and make the one in the workbook the question for the final review. In this way, the anxiety of the reader of the workbook would be reduced, and the results on the final review could still be good, since, by this time, the reader would have read the list of assumptions through twice and responded to the list of possible assumptions which would now be a review question in the workbook.

The third question of the final review gave seven possible decision maker statements or questions and asked the reader to give a brief but specific response that an F-H evaluator would make to each. Four out of eleven field test participants (36%) were competent on this question. The

group, as a whole, did well on three of the seven parts of this question.

The statement, "We want to verify the worth of our program" was suggesting that the decision maker was after public relations data, rather than an F-H evaluation. The correct answer was to include an indication that this was not the real purpose of F-H and that the "worth" of a program should be determined by its goals and the level of accomplishment of those goals. This was answered correctly by seven out of eleven field test participants (64%), and two more received half credit.

The question, "Can't you just tell us what goals we should have?" was answered correctly by seven of the field test participants (64%), and three more received half credit. The correct answer was expected to include an indication that this decision was the proper domain of the decision makers of the program, rather than the evaluator. These two parts of question three were directly related to material presented in the workbook which stressed that public relations was not the purpose of an F-H evaluation, and, also, that the goals of the decision makers are the best guides to the kinds of data they want and will use.

A third part of the question three said "Our main goal is to identify the students' needs," and was answered correctly by nine of the field test participants (82%). The desired response should have mentioned that the decision

makers should really consider whether or not they would prefer a needs analysis to an F-H evaluation. Part of the workbook included a discussion of needs identification as being different from the purpose of F-H which is to provide data about an enterprise for decision making. From these results, no problems with this section could be identified.

The other four parts of question three did cause some problems for the field test group. The question, "What should we do about our program?" was answered correctly by only four of this group (36%). The answer was expected to include that it isn't the role of the evaluator to make this determination, but that of the decision makers of the program. The "wrong" answers, some of which received half credit, most often asked the decision maker what the goals of the program were and if they were being accomplished. In this author's opinion, these answers were not really wrong because they did not have the evaluator presuming to tell the decision makers what to do about their program and they were answers that an F-H evaluator could make. This part of question three did not ask for a specification of the evaluator's and decision maker's roles and should have provided more guidance. In any case, the "Facilitator's Handbook" should be expanded to include other acceptable answers which were not foreseen when it was written.

The question "How can we achieve our goals?" was also written with the expectation of responses similar to the part

just discussed, i.e., a brief clarification of the roles of the evaluator and the decision maker. Five of the field test group (45%) answered it correctly. Of the other six, one received half credit and one had omitted it. The others all suggested that the goals be operationalized or that there be an examination of the activities presently being carried on to accomplish the goals. Neither of these would insure goal achievement, however. By reviewing the assumptions behind F-H and "Other Approaches to Evaluation" in the workbook, the reader should be able to grasp the idea that these decisions should be left completely to the identified decision makers.

The next part of question number three was "We want to make better decisions about our program." This should have suggested that it is not up to the evaluator to help the decision makers make better decisions but that they should try to pick up these skills elsewhere. Only three from the field test group (27%) answered in this way, although six received half credit. The most common alternative answer, which was given half credit, was that the evaluation should be able to help in this endeavor by providing relevant data which has decision maker validity. This is another case in which answers other than those included in the "Facilitator's Handbook" were basically correct. These answers should be included in the "Facilitator's Handbook" for full credit, and the question itself should be expanded so that it is clear to

the reader that it is better decision making skills which are being desired by this decision maker.

The last part of question three was the statement "We don't want to collect that kind of data--it would be embarrassing." Four of the field test group (36%) gave the correct answer which contained the idea that if the decision makers wanted only positive data, they should hire a public relations expert, rather than an evaluator. The wrong answers and those which received half credit varied somewhat for this question. Two persons said that they couldn't collect the data at all. Others said that data weren't value judgments, and, if no one else but the program decision makers used it, why not collect it? Those who answered this question incorrectly should review the section in the booklet on "Other Approaches to Evaluation" which emphasized that the purposes of public relations and evaluation are different, and that, therefore, one should not be undertaken when it is really the other which is desired. The review question relating to this part of the workbook, however, only asked why F-H should not be equated with, among other things, public relations. It should instead ask, "Which of the following questions asked by decision makers indicates a desire for public relations, as opposed to an F-H evaluation?" or something similar to this. In addition, the pertinent question in the final review should be clarified so that it is more

obvious that the decision makers are concerned with their program's image.

The problems that some of the field test participants had with one or more parts of question three, then, either resulted from an unclear statement of the review question itself or could have been clarified by going over more carefully the appropriate sections in the workbook. It did not seem to this author that the difficulties relating to this question had implications for changes in the workbook content, although they did suggest some additions to the "Facilitator's Handbook."

Question number four contained two parts. For each, the reader was asked to select from four choices the one major problem with the evaluation model named. The first section asked for the major problem with the "evaluator-as-expert" model, and the correct choice was the fourth one, i.e., "use of standard evaluation criteria which might not really apply to a given program." All eleven field test participants answered this correctly. The second section asked for the major problem with the "outside evaluator" model, and the correct response was the first one, "use of standardized tests which might be inappropriate for the particular program." Only five participants (45%) answered this part correctly. Of the six who failed to do so, one omitted the question, and one chose the second response, "development of specific data-collection instruments is time-consuming." The

other four chose the fourth response, "provides only public relations type of data." Nowhere was it stated in the booklet that the "outside evaluator" had anything to do with either the development of specific data-collection instruments or with public-relations data. In this author's opinion, the best remedy for the low scores on the second part of question four would be a review of the discussion of other models in the booklet, rather than a change in the booklet's content and/or the question itself. This question was not required for over-all competency in the skill "Defining Evaluation," however, as it was not considered sufficiently important.

Question five was the only other question which was not required for competency in this skill. It asked that the reader write down from memory the major problem of models, vis à vis methodologies, in general. The desired response would have been that models do not contain operational procedures for carrying out an evaluation, while methodologies do. Only three of the field test group (27%) answered this correctly. The booklet did contain the above desired response, but the term "operational" may have been difficult for some readers to understand. A brief example of the descriptive format of the steps in a model as opposed to the prescriptive format of the steps in F-H should be added to the content of the booklet to increase the understanding of this concept.

Question six asked the reader to write down without assistance, a similarity which the "CIPP" and "Discrepancy" models of evaluation share with F-H. The desired answer was that they all focus on decision makers, but other answers given were derived from the workbook content. They were that all three are based on the goals of a program, relate outcomes to goals, or assume evaluation should be on-going and provide continual feed-back. Eight field test participants (73%) answered correctly. There don't seem to be any changes needed here, either in the booklet or in the test question. The other acceptable answers should be included in the "Facilitator's Handbook" in a more specific form.

The seventh question asked the reader to write the three criteria for the success of an F-H evaluation from memory, and, for the second part, to define these criteria. All but one person in the field test group (91%) were able to remember the names of these criteria, but only two (18%) stated the definitions correctly. Many of the incorrect answers substituted "goals" for "decisions" in the definitions of the criteria of completeness and focus. Others confused the criteria of efficiency and completeness, while a few merely stated the definitions poorly, e.g., "is the data provided used in decision making," rather than "the proportion of data provided which is used in making decisions." Two changes are suggested by these results. First, these definitions are rather difficult to understand, especially the transition

from a previous emphasis on goals to an emphasis on decisions. The content of the booklet itself needs to be slightly expanded, perhaps with a review question after the definition of each criterion in order to improve the reader's understanding of these new concepts. Secondly, it is possible that the question itself is inappropriate in its present form. Since students of F-H will learn the three criteria at a later date, asking them to memorize them at this time is a little premature and might cause confusion and frustration when presented along with so much other new material. This question, then, could be much more effective in a multiple-choice format for this introductory booklet.

The final question consisted of three multiple-choice parts. For each one, the name of one or more processes was given, and the reader was asked to check as many of the four following statements as were applicable. There were two right answers for each part, each of which was worth one point. A wrong answer checked counted as a minus point, and, in order to earn competency on this question, a reader was required to earn at least one point (two right answers and a wrong one or one right answer) on each of the three parts. Six of the field test group (55%) were competent on this question. Of those who failed to gain competency, all failed to earn one point on the first part which concerned the goals and parts processes. The most common mistake was that "promote the criterion of focus" was checked and had not been

considered to be a right answer because the booklet linked these processes only with the criterion of efficiency. In fact, however, both these processes include putting lists in priority order, a process which directly affects the criterion of focus. This choice, therefore, should either be eliminated from the question or included in the "Facilitator's Handbook" as a correct answer.

Motivation. The only way in which the issue of student motivation was considered in this field test was by trying to identify participants who were interested in the area of evaluation and, if possible, in F-H. The differences in motivation among the participants is difficult to assess, but the student reactions, which will be discussed later, do suggest that there were differences. If it is true that some of the review questions and some questions in the final review raised anxiety levels, then it is important to try to insure high motivation from the beginning so that the anxiety will not prove too detrimental to the rest of the learning and that individual achievement will be as near competency as possible.

One way of doing this might be a brief explanation, either included in the booklet introduction or given verbally at the beginning by the facilitator, of how the skill "Defining evaluation" was identified. More importantly, it should be emphasized that, on the final list made up from the lists

TABLE III
 "DEFINING EVALUATION": SCORES ON THE FINAL REVIEW AND RELATED VARIABLES

Scores	Special Programs Participants	Hutchinson's Class	"Experts"	Had Eval. Meth. I	Some Experience in Ed'l Eval.	"Familiar with" F-H	Full-time Student	Works Full-time	Other	Competent in the Skill
38		X				X	X			X
37			X	X	X	X	X			
35	X				X	X			X	
34	X				X		X			X
32			X	X	X	X	X			
30			X	X	X	X	X			
30		X			X	X		X		
30	X					X		X		
28			X	X	X	X		X		
28	X				X	X		X		
27	X				X		X			
26		X				X	X			
25	X					X	X			
22		X				X			X	
21		X				X	X			

of three evaluator/methodologists, this skill was number one in priority.

The role of background, training, and experience. Table III indicates the scores of all those taking the final review for this skill. Next to each score is a number of variables such as: whether the person is from the Special Programs group, Dr. Hutchinson's class, or the "expert" group; whether the person had had the Evaluation Methodology I course; whether (s)he had had some experience in any kind of educational evaluation; whether the person considered himself or herself to be "familiar with" F-H; whether the person was a full-time student, worked full-time, or belonged to neither of these two categories; and whether the person had earned competency in the skill. The variables, excluding the group to which each person belonged and whether or not (s)he was competent in the skill, came from the responses to the background questions previously described. The question concerning whether or not the student was "familiar with" F-H was often interpreted as whether the student had ever heard of, or knew anything at all about, F-H. Thus, the results on this question probably only discriminate between those who have never heard of F-H and those who have.

The mean score on the final review for the whole group was 23.7. When the means for the Special Programs group and for Dr. Hutchinson's class were computed separately, they were found to be 29.8 and 27.4, respectively. The mean score

TABLE IV
ANALYSIS OF VARIANCE
GROUP ANALYSIS

Source of Variance	Sum of Squares	Degrees Freedom	Mean Square	F
Groups (Between)	43.8	2	21.9	$F_{J-1, N-J} = \frac{MSS_B}{MSS_W}$
Error (Within)	314.9	12	26.2	$F = \frac{21.9}{26.2}$ $= .83^*$
Total	358.7	14	48.1	

N = 15

J = 3

P > .05*

TABLE V
T-TESTS FOR FIELD TEST OF "DEFINING EVALUATION"

Group Name	N	Mean	Standard Deviation	t	p
Special Programs	6	29.8	4.0	.78	p > .05
"Experts"	4	31.8	3.9		
Hutchinson's Class	5	27.4	6.9	1.13	p > .05
"Experts"	4	31.8	3.9		
Special Programs	6	29.8	4.0	.72	p > .05
Hutchinson's Class	5	27.4	6.9		
"Experts"	4	31.8	3.9	1.06	p > .05
Non-experts	11	28.7	5.3		
Experienced in Educational Eval.	9	31.2	3.5	1.68	p > .05
Not Experienced	6	27.0	6.3		
Full-time Student	9	30.0	5.7	.32	p > .05
Works Full-time	4	29.0	2.2		

for the "experts" was found to be 31.8. It was decided that a one-way analysis of variance would be done for the three groups, even though the groups were not randomly chosen from different populations. The purpose was not to generalize any results by using this technique, but only to provide an index of the size of the difference in the means of these three groups for this particular field test. In addition, it was determined by the Bartlett test (Winer, 1962, p. 95) that the homogeneity of variance assumption necessary to do an analysis of variance was violated. This violation would have affected the interpretation of the results, had significance been found, but does not alter the fact that the differences here were not statistically significant.

A one-way analysis of variance ($p > .05$) indicated that these means were not statistically different (Table IV). In addition, t-tests showed that the group means did not differ significantly from each other, nor did the "expert" group differ significantly from the other two groups combined which had a mean of 28.7. It can be seen by Table III that this is the same as saying that there were no significant differences in the means of the group that had taken the Evaluation Methodology I class and those who had not. It was also found that the means of the group which had had some experience in educational evaluation (31.2) and the group which had not (27) did not differ significantly, nor did the means of the full-time students (30) differ significantly from those who worked full time (29).

It can be said, then, that there were no variables

identified which related to significant differences in scores for those individuals tested. It was especially interesting that the "experts" did not do significantly better than the non-expert group which suggests that those who went through the booklet only once were able to do nearly as well on the final review as those who had had prior training in the methodology and experience in the field. This result only applies to this particular field test, however. This booklet is to be revised on the basis of the discussion above and the further discussion which follows. It is hoped that those who use the revised booklet, and who also go back over those materials which caused difficulty, would do even better than an "expert" group which had not been trained in this skill in this particular way.

For future field tests of revised versions of these materials, if groups similar to the ones used here could be identified, it would probably be unnecessary to apply the background questions again. As formulated here, they did not furnish variables which related to significant differences in scores achieved on the final review.

Student reaction. Of the eleven field test participants who responded to the "Student Reaction to the Handbook 'Defining Evaluation'" questionnaire, six gave the booklet a "1" on the five-point scale "easy to follow" to "difficult to understand," thereby indicating they thought it very easy to follow. Four gave it a "2" on this scale, and there was one "3." On the scale "clearly stated" to "confusing or jargony," there was more variation in responses. The results were as follows:

three "1's"; four "2's"; one "3-4"; two "4's" and one "5." All but two participants rated the booklet "1" on "was comfortably completed within the given time" to "took too long." These two gave it a "2" on this variable.

With respect to the review questions, the question which asked the reader to check off on a list of descriptive phrases as many as were applicable or to fill in a phrase under "other," nine persons checked "appropriate to the workbook content." Three check "challenging," two checked "too easy," and two wrote in comments, one of which said that the multiple choice was too easy but the fill-ins were challenging; and the other that the questions were too simplistic because they asked for rote memory, rather than testing understanding.

The responses to the same question but in reference to the final review yielded eight checks for "appropriate to the workbook content," five checks for "challenging," and one "too difficult." Three persons made additional comments. One said that it was a bit easy because of the number of multiple-choice questions, and another wrote that it didn't really test understanding. A third said that she'd have liked to have had more fun doing it.

Other responses to the workbook were given in an open-ended question at the end of the questionnaire. Five out of six of the Special Programs participants liked the review question format, with the opportunity for immediate feedback, better than the final review. A few were dissatisfied with

the amount of responses on both the review questions and the final review which called for the reader to recall an answer from memory. Some felt that the total emphasis should be more on understanding than on memory, and a few disliked the use of "jargon." This was interpreted by this author to mean the inclusion of specific names of other evaluation models, their authors, and terms specifically related to research and evaluation, e.g., "variable," "post hoc data," "control group," "random," etc. Related to this comment was the feeling by some of the Special Programs participants that these names and terms were distracting because the reader did not know what (s)he was going to be asked to remember. They felt that a specific list of objectives in the beginning would have helped them to know on which concepts they should concentrate.

For those learning about F-H solely through workbooks, the comment about "jargon" is a valid one. It is difficult to proceed through self-instructional materials when there are a number of names, words, or phrases which are unfamiliar or which seem to be included for effect, alone. The serious student of F-H will be involved in classroom instruction during which these terms should be clarified. The important thing is for the F-H evaluator not to appear to be using "jargon" when working with decision makers. This doesn't mean that the evaluator doesn't need to understand and be able to use the "jargon" for his or her own benefit

and where these terms are common among professional evaluators.

Another person commented that the order of the presentation of topics in the workbook was wrong and that the old assumptions people have about evaluation should be presented first so that they may be led more gradually into the newer F-H approach. Other suggestions included the use of case studies to illustrate other evaluation models and examples of desired responses for question three of the final review which was the one with the seven decision maker statements to which the reader, in the role of an F-H evaluator, was asked to respond. Finally, one participant also added that she would be interested in knowing the retention of this material over time.

These comments make some very practical suggestions for the improvement of the workbook. The most important, in this author's opinion, concerns a reexamination of what learning is really desired, for users of these materials. This will involve a careful consideration of the kinds of questions asked within the booklet and as a final review. The suggestion of including the objectives for the reader in the beginning may well be helpful in promoting learning, and the idea of including more examples, both within the content of the booklet and as an explanatory device for the questions on the final review, should have its place in the revised form of "Defining evaluation." In addition, the order of the

presentation of topics should be reexamined, and the possibility of making the booklet's style somewhat lighter should be considered. The problem that some participants had with questions asking for recall of a term, idea, or set of ideas, is more of a motivational problem than an implicit fault of these materials. An evaluator will need to have some of this information at hand, and, if the participant doesn't want to learn it, perhaps he/she is not really interested in becoming an F-H evaluator.

Other recommendations. Other areas that might be investigated when the booklet is to be revised are whether there were too many instructional objectives for one booklet, whether any were inappropriate, and whether the competency requirements should be modified in any way. The two objectives which should be deleted or changed are: that the student should be able to recognize the major problems behind some other evaluation models; and that the student should be able to give the major disadvantage of models, as compared to methodologies. These two objectives could conceivably be the basis for two full courses of study. Since merely mentioning the names of some models with a one-sentence description of each does little to furnish the student with a sound understanding of what models are all about, and might add substantially to the student's confusion, it would seem that these objectives should be substituted by one that stresses that there are different definitions of evaluation in use in the

field. Also, since the questions on the final review which relate to these objectives were not considered necessary for overall competency in the skill, it would not be difficult to delete these objectives.

The results of the final review and the student reactions do not suggest that there was an attempt to cover too much material. This would not indicate, therefore, that there is a need to reduce the number of instructional objectives for this reason alone.

In this author's opinion, the competency requirements for this skill, with the exception of that for question number seven discussed above under "Final review," were appropriate. The fact that only two persons gained competency after the first round indicates that the material needs to be studied, and that most readers should expect to cycle through the materials more than once, as directed, before they become competent. The aim of the workbook is that the reader gain competency on this skill, however, and the fact that only two did so on the first round is a further indication that revisions are needed.

The necessity of the workbook. Of the four "experts" who took the final review, none were competent in this skill. Two missed one question, and the other two missed two, although the questions asking for the F-H definition of evaluation and a similarity among the "CIPP" and "Discrepancy" models and F-H were not missed by anyone in this group.

These results indicate that the "experts" can also benefit from the use of the booklet, although they would probably only need to go through it once to achieve competency on this skill, because of their prior training and experience.

Final comparisons. The question which yielded the most interesting results for this group, as opposed to the field test group, was question three, which contained seven questions or statements which might be made by decision makers to an F-H evaluator. The two questions "What should we do about our program?" and "How can we achieve our goals?" were answered correctly by three out of four "experts" and by the whole group, respectively. They did much better than the field test group because of their experience with real decision makers in the course of their evaluations during which it was likely that they had to answer very similar questions.

The statement "Our main goal is to identify the students needs" was only answered correctly by one "expert," although 82% of the field test group got it right. Two "experts" summarized the goals process and suggested that the decision makers go through it, while the other merely agreed to accept this as the main goal. These are possible responses to this statement but do not seem to be the best ones, suggesting that the "experts" might benefit from the section in the workbook, "Other Approaches to Evaluation."

Finally, the statement, "We want to make better decisions about our program" was answered correctly by only one

of the "experts." The field test group also did not do well on this part, with only 27% answering correctly. As mentioned in the discussion of the final review above, there seems to have been some ambiguity in the question itself, as well as some correct answers being omitted from the "Facilitator's Handbook."

It seems, then, that the question concerning needs which gave the "experts" difficulty would have been answered correctly, had they gone through the workbook materials. They would also profit by using the revised version of the workbook for the decision maker statement, "We want to make better decisions about our program," in addition to the other questions shown in Table II and Table VI for which competency was not achieved. Table IX summarizes the results and their implications for the field tests of both workbooks.

Design of the Field Test for "Dealing with a Lack of Decision Maker Cooperation"

The second workbook developed during the course of the work for this dissertation was called "Evaluation Handbook #2, Dealing with a Lack of Decision Maker Cooperation." It was originally conceived as an aid for those who had been trained in F-H but had not yet had any field experience doing complete evaluations. Some of the scenarios contain references to major steps in F-H, and knowledge of the methodology as a whole would seem to be helpful for one's understanding of how this lack of cooperation might jeopardize the entire

evaluation. Again, a small group was desirable so that the major problems could be identified.

The individuals contacted, therefore, were ones who had had the "Evaluation Methodology I" course taught at the School of Education, University of Massachusetts at Amherst, by either Dr. Thomas Hutchinson or Dr. William Gorth. Three people agreed to go through the booklets, take the final review, and react to both during a two-hour session on March 12, 1975. It was estimated that the completion of the booklet would take about an hour and a half; the final review, about half an hour; and a few minutes for the background and reaction questions. Because the time allotted for field testing for this dissertation was rapidly coming to a close and scheduling busy graduate students for the same two-hour time period was becoming very difficult, it was decided to allow any who were willing to go through the materials at home. They were instructed not to look at the final review before going through the booklet; to go through the booklet completing the review questions as they occurred; to complete the final review without reference to the booklet; and to respond to the background questions and student reaction questionnaire afterwards. Three more persons agreed to do this and returned the materials by March 21, 1975.

Finally, it was suggested by Dr. Hutchinson that perhaps it was not really necessary for one to have had the prerequisite course in order to successfully complete the second

workbook and that his class, which had participated in the field test of the first handbook, "Defining evaluation," might be invited to attempt the second one as well. This was then suggested to his five students, but none of them returned the materials in time for the results to be included in the following discussion. Thus, six persons made up this field test group.

Background questions and student reaction. The same questions were used for eliciting background information and students' comments about this workbook, the review questions, and the final review as were used for the previous workbook. Again, no pre-test was designed. Not only would a pre-test not fit the purpose of this field test, but it was assumed that even students of F-H would not be competent in this skill since this content material is not currently being taught in the methodology courses and is even difficult to learn through limited experience in the field. The background questions were to be used to determine whether differences in scores might be attributable to differences in background or experience in educational evaluation. The questions on student reaction were to be used for revising the booklet so that it might better achieve its instructional objectives.

Experts. The same four "experts" who took the final review for the first skill, "Defining evaluation," also took the final review for this skill, again without going through the booklet. Because this skill is of a different nature than

the first and is a more difficult skill to learn, it was not necessarily assumed that the "experts" would do as well or better than the field test participants. Since this skill had ranked so highly on the final list, however, it was considered desirable to find out how some of those who had received all available training in F-H as well as having participated in or conducted full-scale F-H evaluations would do on a test of this skill. The completed final reviews were returned to this author by March 21, 1975.

Results and Discussion of the Field Test of "Dealing with a Lack of Decision Maker Cooperation"

The completion of the workbook took about an hour and a half, and the final review took from twenty to thirty minutes. This was approximately the anticipated time for each. No questions were asked when the first group of three went through the booklet and the final review. No one in either group gained competency in the skill as a whole.

Review questions. Of the eleven review questions distributed throughout this handbook, five were answered correctly by all six field test participants. These included the first question which referred to some general information in the introduction, and four questions which asked the reader to fill in ways to avoid and/or alleviate the following problems: the decision maker has philosophical disagreements with F-H; the decision maker is not "methodologically

oriented"; the documentation is dull, dry, or threatening; and the decision maker was wrongfully identified by the contract decision maker. In addition, for the second problem mentioned, the review question included a description of the decision maker behaviors and asked the reader to fill in the problem name. The documentation problem also asked the reader to underline one answer to the following question: "The problem of dull documentation is (difficult, somewhat difficult, fairly easy) to avoid."

These four examples of lacks of decision maker cooperation caused fewer problems for the field test participants, possibly because these problems were defined more clearly than others. They seem to be simpler with fewer "symptoms" which could be appropriate to more than one problem, and, perhaps for this reason, the courses of action open to the evaluator were easier to remember.

Only three review questions were missed, either partially or completely, by more than two of the field test group. These questions referred to the following problems: the decision maker doesn't want to know what (s)he is doing wrong: the decision maker has a "status hang-up"; and the decision maker is insecure and/or under pressure. In one case part of an answer was omitted, and two others answered the first part of one question incorrectly. All other responses to these three questions were incorrect, however, because a course of action which had been designated as a

measure to prevent the problem ("avoid") was included as an answer to a way by which the evaluator could help solve ("alleviate") it, and vice versa. This confusion occurred for other review questions and occasionally on the final review as well. The differentiation between suggested actions for the prevention and alleviation of these problems was often not well-defined, and sometimes they were very similar except for the recommended time of application during the course of the evaluation. Also, these three problems had symptoms which were fairly similar which made it very difficult for users of the handbook to commit the solutions to memory the first time through.

Changes should be made in the formats of the review questions, especially for those problems which involve similar decision maker behaviors. It must become boring, as well as confusing, to read through so many scenarios, memorize so many procedures for the prevention and alleviation of these problems, and then have to respond in practically the same way for each review question. A variation of formats could help increase interest and motivation, as well as improving achievement.

Final review. The first question of the final review consisted of three scenarios for each of which the reader was asked to give the following four answers: the most likely problem causing the lack of decision maker cooperation; one way of avoiding the problem; one way of alleviating the

TABLE VI
 "DEALING WITH A LACK OF DECISION MAKER COOPERATION": FINAL REVIEW QUESTION
 COMPETENCY ACHIEVED BY THE FIELD TEST GROUP AND THE "EXPERTS"

Final Review Question #	Field Test Group						"Experts"			
	#1	#2	#3	#4	#5	#6	#1	#2	#3	#4
1	X	X								
2		X	X		X		X	X	X	
3	X	X	X	X	X	X	X	X		
4						X			X	
5	X	X	X	X	X	X	X			X
6		X	X		X					

problem; and another possible problem behind these decision maker behaviors. The three scenarios concerned the problems of a decision maker not wanting to know what (s)he is doing wrong; the decision maker who has a "status hang-up"; and the decision maker who was wrongfully identified by the contract decision maker. Only two from the field test group (33%) gained competency on this skill, as is shown in Table VI. Competency required a minimum of eight, out of a possible twelve, correct answers.

The field test group averaged six and one-third points out of twelve on this question. There were no dramatic trends found when the responses to question one were analyzed. In general, there were slightly fewer errors made on the second part of question one (decision maker has a "status hang-up"), and slightly more made on the third part (the decision maker was wrongfully identified by the contract decision maker).

When the responses were analyzed in terms of the four required answers to each of the three parts, of the eighteen possible responses (six participants and three scenarios), the field test group achieved 72% correct for the probable problem diagnosis. There were similar trends for the parts asking for ways by which the evaluator could avoid and alleviate the problem. When the reader was asked to give another possible problem which might be causing the given set of decision maker behaviors, the field test group only achieved

28% correct answers. Two of this group, either accidentally or intentionally, had omitted this fourth part for all three scenarios. In addition, the section concerning other possible problems had been added as an after-thought to each of the ten problem discussions in the booklet, and none of the review questions asked anything about this section at all. The field test participants, therefore, probably tended to ignore this discussion. This weakness in the booklet should be remedied by improving the review questions, as well as by trying to insure that the other possible problems stem logically from the set of symptoms already given.

The second question asked the reader to match seven problem names with the most likely set of symptoms, out of nine given. There were thus two sets of symptoms with no match. As is shown in Table VI, three in the field test group (50%) were competent on this question.

One set of symptoms was matched incorrectly by two-thirds of the field test group. This set was "wants you to identify program goals; has difficulty understanding the decision maker's role; often asks evaluator's opinions; expects to have little direct involvement in the evaluation." It was intended to be one of the sets without a match, but two persons linked it with "decision maker was wrongfully identified by the contract decision maker" one with "decision maker is not 'methodologically oriented'"; and one with "decision maker has a philosophical disagreement with F-H."

These symptoms were originally taken from the problem "the decision maker has a lack of understanding of F-H." It is evident that this set of symptoms could be considered to fit with any of the problems with which they were matched. Thus, the question of whether or not these four problems are really discrete or whether a sufficient number of appropriate symptoms can be identified so that they become discrete should be considered. It may be that all or some of them may be found to be overlapping to a degree so that the problems should be diagnosed in a more general way. It would seem that the problem of a decision maker who has been wrongfully identified could overlap with any of the other problems mentioned or with none of them. The other three problems seem to be separate entities, but a reexamination of all of them should be undertaken in order to minimize the reader's confusion and to insure that these materials will furnish valid information which will be usable in practical settings with real decision makers.

The third and fourth questions required the reader to respond with the correct problem name. The first asked for the most common problem which is most often used as an excuse by decision makers to disguise other reasons for a lack of cooperation. The correct answer was a lack of time, and was given by everyone in the field test group. The only other question with this good a response record was number five. Question four asked which problem was the most difficult to

diagnose and the most likely to underlie several other problems. The correct answer, that the decision maker was insecure and/or under pressure, was given by only one of the field test participants.

By going back through the workbook itself, it becomes clear why two such similar questions received such different response patterns. The problem of a decision maker having a lack of time was specifically cited as being the most common problem and the one most likely to be used to cover up other problems. The problem of decision maker insecurity, however, was never referred to in the booklet as the most difficult to diagnose or the most likely to underlie other problems. It was termed "common" and "difficult to pin down," as well as being likely to cause other problems. In the summary of general procedures at the end, it was stated that this problem may not become apparent until the third problem-solving attempt, but that was all that was said about it. In addition, the problem of decision maker insecurity has more similarities to other problems, e.g., the decision maker not wanting to know what (s)he is doing wrong; the decision maker having a "status hang-up"; and the decision maker having been wrongfully identified by the contract decision maker, than does the problem of a lack of time. This makes the former is more likely to be confused with other problems than the latter. Of the field test group, two answered this question with the problem of the decision maker having a

"status hang-up," and two said it was the decision maker not wanting to know what (s)he was doing wrong.

This question could be improved by providing the necessary information in the booklet. It would also be beneficial to provide a diagram of the hierarchical nature of the problems presented with those which are the easiest to recognize on the top and those that are the most difficult to diagnose and the most likely to cause other problems on the bottom. The latter problems could be linked by arrows to those which they may cause or with which they would probably overlap.

The entire group was competent for the fifth question, as for the second. It asked the reader to give seven things an evaluator can do in the beginning of an evaluation to avoid problems due to a lack of decision maker cooperation. This should have been fairly easy for the field test group because the prevention of each of the ten problems was discussed in the booklet, and a summary of eighteen courses of action open to the evaluator was given at the end of the booklet. It appears from these results that modification of the content or the question is not required at this time.

The sixth and last question of the final review asked the reader to give, in order, four general steps that an evaluator might take when faced with a lack of decision maker cooperation. The three out of four points required for competency (one for each step given) was achieved by half of the field test group. A six-step general problem-solving

procedure was outlined in the last pages of the booklet, and this must have been used by the three who were competent on this question. All would probably have benefited, however, by more of a discussion of this general approach, since it was not mentioned in the booklet until the end, except for the section on "Other possible problems" which, as was previously mentioned, was not tested at all by the review questions. Increasing the content concerning this subject and providing appropriate review questions would be valuable, especially since some of the distinctions among specific problems are tenuous. It is, therefore, difficult to remember the specific suggested activities, and, in that case, an evaluator should have at his or her disposal a general method of attacking the problem and finding solutions.

Motivation. For this workbook, as for the previous one, it was hoped that the student's interest in the area of evaluation would provide the needed motivation to go through these materials in a conscientious way. The fact that all the field test participants were full-time students who had had the Evaluation Methodology I course was considered to be evidence of this needed interest. Again, however, student reaction to this workbook and the final review indicate differing levels of motivation.

The reactions, which will be discussed later, included the widespread feeling that the materials took too long. This, coupled with the uniformity of format of the review

questions already discussed, could be two factors which might act to lower the student's interest in the whole process. The possibility of dividing the booklet into two sections, each of which would contain five problems concerning a lack of decision maker cooperation and could be administered separately, might be considered. Other alternatives might include: spreading the instruction over several sessions; using the workbook along with supplemental course work, and giving the final review at the end of the course; or just building in some discussion time in the middle of the course of the workbook and at the end, prior to the final review. In this author's opinion, the second suggestion is the best because of the quantity of material covered. This would also allow for more opportunities for role-playing situations, thereby adding to the students' expertise in this skill. These are important considerations, especially since the students, as well as the "experts," recognized the value of the content material, and it would be unfortunate if the style and structure of the booklet itself acted to impede learning.

Again, as with the previous booklet, it might also help to inform users of the materials, either as a part of the introduction to the booklet or through a verbal introduction given by the facilitator, of the importance of this particular skill to the training of F-H evaluators. It could help increase motivation for them to know that this skill was located on the second priority level of the over-all skills

TABLE VII
 "DEALING WITH A LACK OF DECISION MAKER COOPERATION": SCORES
 ON THE FINAL REVIEW AND RELATED VARIABLES

Scores	Field Test Partici- pants	"Experts"	Had Eval. Meth. I	Some Exper. in Edu. Eval.	"Familiar with" F-H	Full-time Student	Works Full-time
26	X		X	X	X	X	
23	X		X		X	X	
22	X		X	X	X	X	
19	X		X	X	X	X	
19	X		X	X	X	X	
18		X	X	X	X	X	
17		X	X	X	X	X	
16		X	X	X	X		X
12	X		X	X	X	X	
5		X	X	X	X	X	

TABLE VIII
 T-TEST FOR FIELD TEST OF "DEALING WITH A LACK
 OF DECISION MAKER COOPERATION"

Group Name	N	Mean	Standard Deviation	t	p
Field Test Group	6	20.2	4.8	1.82	$p > .05$
"Experts"	4	14.0	6.1		

list which was a combination of priorities obtained from this author, Dr. Hutchinson, and Dr. Benedict.

The role of background, training, and experience. As Table VII shows, the only reasonable way in which those using the materials for "Dealing with a lack of decision maker cooperation" could be divided on the basis of the variables given was as field test participants and "experts." The other variables resulted from the responses to the background questions given to the field test participants and this author's knowledge of the background of the "experts." Because these participants as a whole had had more experience in the area of educational evaluation than the groups tested for the previous booklet, the background questions might have included some different variables, such as whether or not the participants had had a statistics or research course or whether or not they had ever had any difficulties with a decision maker while engaged in carrying out an F-H evaluation.

The means of the scores for the field test group and the "experts" were 20.2 and 14.0 respectively. A t-test was done on the means from the two groups in order to provide an indication of the size of this difference. Again, the intention was not to generalize these results to other groups. The Bartlett test (Winer, 1962, p. 95) indicated that the assumption of homogeneity of variance was not violated. These were shown ($t = 1.82, p > .05$) not to be significantly different by the application of a t-test. It was expected, however, that the "experts" would do worse than the field test

group, since the material in the booklet was not covered at all when the former received their training, and it is very difficult to become competent in this skill just by having had experience using F-H in practical settings. If anything, the scores of the "experts" were slightly inflated, since one had worked with this author on the initial attempt to define this skill, and the other three had all served as tests of completeness during this defining process.

Again, these results hold only for this field test, and, since the handbook is to be revised on the basis of the results discussed here and the student reactions which follow, they are particularly short-lived. With the revised edition of the handbook and given the opportunity to go back over content material which they failed to master during the final review, the field test group should be able to do significantly better than a traditionally-trained group of "experts" who had not participated in the definition of the skill.

Student reaction. On the scale "easy to follow" to "difficult to understand," three of the six field test participants gave this workbook a "1," one a "2," one a "3," and one a "4." This shows mixed feelings, but only one person ranked the booklet on the "difficult" end of the scale. On the scale "clearly stated," to "confusing or jargony," the split was very similar. There were two "1's," one "2," two "3's," and one "4." Again, only one person (the same one)

rated the workbook on the "confusing" end of the scale. A more interesting split came on the responses to the scale "comfortably completed" to "took too long." Three gave the workbook a "1" on this variable, two gave it a "4," and one a "5." This shows that half of the field test group thought the booklet was manageable within the allotted time, while the other half thought it was too long. It is interesting that the three who did not feel the materials took too much time were those who went through them at the same time in the same room, while the three who did the workbook, final review, and background and reaction questions on their own at home felt that they were too time-consuming. It should be noted that the intended use of this workbook would take place in a classroom setting.

On the topic of the review questions, there were four checks next to "appropriate to workbook content," three next to "challenging," and one next to "too difficult." The latter was from the same person who felt the review questions were too difficult. Other comments given were that the expected format of the final review was multiple choice questions or something other than "total recall." A few wrote that the summaries of suggested actions for the evaluator were good and clear.

Other reactions to the materials as a whole followed no pattern. One person gave careful, detailed suggestions about specific structural problems which he felt increased the

reader's confusion. Others commented that certain courses of action seemed very sound, while the suggestion that the evaluator actually give his/her opinions, if they are asked for repeatedly, was frowned upon by at least two field test participants. One felt the written reinforcement after the review questions was not necessary, but that the suggestions for reviewing certain pages were helpful. Two other comments he made were that a page explaining the purpose behind the booklet, its instructional objectives, and the expected levels of user competency was needed, and that the reader should be cautioned that all problems involving decision makers can't be anticipated.

Another problem which a couple of the participants commented upon was discussed above under "Final review," i.e., that of the confusion between the sections describing ways the evaluator could anticipate a problem and ways (s)he could help alleviate it. One said that often the same activity serves both purposes and that it's the time of application which makes the real difference. Another objected to the occasional use of "help solve" in place of "alleviate," and "anticipate" for "avoid." Finally, one participant wrote that he suspected that the identification of such problems would be much more difficult in an "on-the-job" situation and that the problems with decision makers are really multifaceted and difficult to identify through a straight "cause-effect-solution" approach.

The main problems that the field test participants seemed to emphasize, then, were the length of time it took to complete the materials and the need for a clearer distinction between actions to be taken to "avoid" the problem and to "alleviate" it. Also, a few additional courses of action were suggested for various problems, and some consideration should be given to changing the formats of the review questions to help make them a bit clearer.

Other recommendations. It is encouraging that the problems that were identified through this field test did not include the style of the booklet being unclear or "jargony," as was the case with the previous workbook. The most important step to be taken to revise these materials, in this author's opinion, would be the further clarification of the problems themselves. There was some degree of overlap of "symptoms" and solutions in at least half of the problems in the booklet. These were: the decision maker not wanting to know what (s)he is doing wrong; the decision maker having a lack of understanding of the methodology; the decision maker with a "status hang-up"; the decision maker who is insecure and/or under pressure; and the decision maker who was wrongfully identified by the contract decision maker. Further definition of these problems, including an identification of the behaviors which are specific to each problem and those which are common to two or more, would be most helpful to those who will need to use this material in the course of

TABLE IX
RESULTS AND IMPLICATIONS FOR REVISIONS FOR BOTH SETS OF MATERIALS

Possible Implications	Sources of Results					
	"Defining Evaluation"			"Dealing with a Lack of Decision Maker Cooperation"		
	Review Questions	Final Review	Student Comments	Review Questions	Final Review	Student Comments
Expand "Facilitator's Handbook" to include more answers		#3, #6, #7				
Expand question itself to include an example		#3	More examples for final review Q. #3			
Expand related review question		#3				
Add a review question		#7			#1, #6	
Change related question on final review	#5					
Change format of question itself		#7		Review questions in gen. for variety		

TABLE IX--Continued

Possible Implications	Sources of Results					
	"Defining Evaluation"			"Dealing with a Lack of Decision Maker Cooperation"		
	Review Questions	Final Review	Student Comments	Review Questions	Final Review	Student Comments
Expand related workbook content					#1, #2	"alleviate"- "avoid" confusion
Expand related workbook content		#5	more examples needed		#3, #4, #6	
Reexamine competency criteria for this question	#2, #5		where memory required			
Include objectives in the beginning of the workbook			for clarity			
Change instructional objectives		#4, #5				
Improve workbook structure		in general				too long

doing evaluations. It would also provide the basis for improved review questions and help clarify the best ways to keep the problem from occurring, and those which would be helpful in alleviating it. It would hopefully become clearer which behaviors of decision makers constituted problems in themselves and which ones were symptomatic of deeper problems. The suggested actions for the evaluator could also be expanded, possibly with a rationale explaining them and discussing why some other possible actions might be inappropriate.

It may be discovered, however, that some of the designated problems are not really separable, at all, i.e., that they make up a greater problem. Methodological research would have to be done to validate the decision maker behaviors that go with each problem, and new problems will have to be formulated when the old ones are shown to be inaccurate. In addition, techniques should be developed for dealing with the situation in which there is more than one problem operating at the same time with the same decision maker.

The necessity of the workbook. There was no question in this final review which was answered correctly by everyone in the "expert" group. In addition, there were two questions, the first one with the three scenarios and the last one involving the general procedures for dealing with a lack of decision maker cooperation, which were answered incorrectly by the entire group. The two questions which the whole field

test group got right were each answered correctly by only two of the "experts." These results indicate that this workbook is a necessary addition to existing F-H training procedures, and that even experience in the field is not sufficient to guarantee competency in this skill.

Further comparisons. For the first question involving the three scenarios for which none of the "experts" was competent, only three of the field test group achieved competency. The experts, however, consistently achieved approximately 42% correct (four "experts" and three scenarios) on all four answers required. This was worse than the field test group on the first three sections--probable problem diagnosis, ways to avoid the problem, and ways to alleviate the problem--and better on the last section--other possible problems. The first three trends were expected because of the lack of training in this skill given to the "experts." For the fourth part, due to the problem connected with the "other possible problems" section in the workbook discussed above, the field test group only achieved 28% correct answers. The "experts," however, who had all had experience using F-H during evaluations, were able to answer correctly more often because of this.

The last question asked the reader to list four general steps that an evaluator could take to deal with a lack of decision maker cooperation, and was missed by every "expert." Three of the field test group were competent on this question,

however. If the suggestions discussed above for this question are included in the revised booklet, the "experts" could definitely benefit from them, since, despite their experience, they don't seem to have developed general methods of solving problems concerning decision maker cooperation.

The third and fifth questions which were answered correctly by the entire field test group were only answered correctly by two of the "experts." These were the questions which asked which problem was the easiest to diagnose and most likely to be used to disguise other ones, and what were seven things that an evaluator could do to anticipate a lack of decision maker cooperation. Although only two "experts" missed each of these, they could evidently benefit from using the workbook. Very often, some of the evaluator actions they listed in answer to the fifth question were more appropriately actions that an evaluator could take to alleviate one or more of the problems, rather than to avoid them. This distinction might become clearer through the use of the workbook.

CHAPTER VI
SUMMARY, CONCLUSIONS, AND IMPLICATIONS
FOR FURTHER RESEARCH

Summary. The need for sound evaluation practices in the field of education is well known (Scriven, 1967; Glass, 1969). Because Fortune-Hutchinson (F-H) is a methodology, rather than a model, it offers the user a standardized, systematic, operational set of rules and procedures for evaluation which has as its purpose the providing of data for decision making. It seemed to this author, however, that becoming proficient in the use of F-H has not been a well-defined task to date. It involved learning and understanding the major steps and applying them in an evaluation situation outside of the classroom, but it was not clear when or how a person would know that he or she had become a skilled F-H evaluator. This work was undertaken, therefore, in order that this achievement could be defined. The following tasks were deemed necessary to meet this need: a systematic identification of the skills an F-H evaluator needs to have; a break-down of these skills where necessary into well-defined skills that can be taught and for which achievement can be tested; the development of instruction for teaching, and instruments for determining the achievement of, these skills; and the field testing of the instructional materials and

instruments. The purpose of the field test would be to identify any major problems with the materials.

A comprehensive skills list was developed from a variety of sources, especially the latest documentation of F-H; a brief questionnaire distributed to experienced students of F-H; related dissertations; and other pertinent literature. The list was put in order by means of the criterion "importance for being a successful F-H evaluator," and the final list combined the ordered lists of this author, Dr. Hutchinson, and Dr. Benedict. This list is included in Chapter III.

The two most important skills which were chosen for the development of instructional materials were "The evaluator should be able to define evaluation" and "The evaluator should be able to deal with a lack of decision maker cooperation." The latter skill had to be further defined before the development of instructional materials could take place.

Self-instructional workbooks and criterion tests based on instructional objectives were designed and field tested for both skills. The workbook for "Defining evaluation" (see Appendix A) was field tested with individuals who had in some way demonstrated an interest in educational evaluation, and the field test group for "Dealing with a lack of decision maker cooperation" (see Appendix B) was made up of individuals who had had a course in evaluation methodology. In addition, both groups responded to questions about their

backgrounds and reactions to the instructional materials. These are included in Appendix C.

A group of four "experts" who had received all available classroom training in F-H and each of whom had conducted an evaluation using F-H were asked to take the two criterion measures without going through the workbooks. The purpose of this was to insure that the materials were really necessary additions to the existing training procedures.

The results of both field tests were analyzed with respect to their implications for improving the workbooks. Two review questions from the first workbook were found to be inappropriate and to require change. More clarification was needed for two questions on the final review, along with a reexamination of the competency criteria for a third. Also, two instructional objectives concerning other evaluation models were found to be inappropriate for the purpose of the workbook, and it was suggested that a single objective replace these two. These suggestions would also have implications for changes in the content of the workbooks.

For the second workbook, the major problem identified by the field test group was a confusion between suggested courses of action an evaluator should take to avoid certain problems and those which would alleviate them. Some of the problems did not seem to be separate entities but overlapped with other problems so that it was difficult to identify them, much less to plan appropriate actions. Also, some of

the members of the field test group felt the materials took too long, and suggestions were made to alter the formats of the workbooks.

Conclusions. The skills list appears to be as complete as possible at this time. It is comprehensive, covering six categories of skills, and provides the trainer of evaluators with new directions, such as "Interpersonal skills," not previously undertaken during F-H instruction. The skill-based approach to training evaluators has the advantage of being more uniform and more complete than the present system. It is possible that some of the problems that F-H evaluators now face in the field can be handled in advance in the classroom, since many more specific issues may be addressed through this more systematic training approach.

The self-instructional approach to some of the skills needed by F-H evaluators would be a worthwhile counterpart to regular classroom instruction. These materials have been shown to be usable, only requiring between one and two hours of time. Their use can be varied with the addition of extra classroom activities such as discussions and simulations, if desired.

The fact that none of the four "experts" was competent in either skill indicates the necessity of these materials and the skill-based approach for training evaluators. The question of whether or not their prior training was carried out well enough is difficult to determine. It is known that this

training did not include the material covered in the second workbook which explains their poor showing on the related final review. For the first workbook, they did fairly well no one missing more than two of the six questions required for competency. In this author's opinion, their lack of competency on the related final review in this instance was more a reflection of their having forgotten material originally learned than of having had poor training in this area, since their original training had included most of this material.

In addition to providing opportunities for improved training of F-H evaluators, these materials can also be used by persons who do not necessarily want to become proficient in F-H but who would like to pick up some of the necessary skills. The Special Programs group who had not had any classroom exposure to F-H demonstrated that they could do almost as well as the "experts" on the final review for "Defining evaluation" (their mean was 29.8 as opposed to 31.8 for the "experts"). Thus, these materials do not require a classroom setting for their use. Skill-based materials can allow individuals who desire less complete instruction in F-H to select the skills which are most suitable to their needs and to use the related instructional materials. This has implications for making the use of systematic evaluation procedures more widespread.

Thus, this study has applied some well-known procedures for planning curricula to the area of training for F-H

evaluators in order to make this practice more complete and appropriate to students' needs. It has indicated one way in which the skills list might be used in the future, but many variations will no doubt become evident through the efforts of others interested in training evaluators at many levels.

Implications for further research. As these materials are revised and field tested again, an attempt should be made to determine the reliability and validity of the final reviews. In order to be reliable, the final reviews should measure the achievement of the competency criteria consistently over time. A field test of each set of the revised materials could be set up to allow for three administrations of the final review--one immediately after participants had gone through the workbook, one about a week later, and the third about three months later. The last testing would help to determine whether or not there was a loss of competency by question and overall which was due to a lack of retention of the material. The differences in competency earned from the first to the second administrations would be caused by a lack of reliability of the instrument and, if the second testing shows a decrease in the number of participants earning competency on any of the questions, possibly by a lack of retention. The actual lack of retention cannot be known after the second testing, but the percent of reliability for each question could be computed by counting the number of responses which either earned or failed to earn competency on the

question for both testings and dividing this by the number of participants. The overall percent of reliability would be found by dividing the total number of "consistencies" by the product of the number of questions with the number of participants, where a "consistency" would be defined as a response to a question by an individual which either earned or failed to earn competency for both administrations of the test. A percent of reliability which is judged to be too low would imply that changes would be required in the questions of the final review so that the possibility of a participant answering questions correctly solely by guessing would be diminished.

In addition, the stability of the behaviors being measured also affects the reliability of the test. All that is known about this now is that the behaviors of the "experts" which overlapped with those elicited by the final reviews were largely a result of training and field experiences which took place two years ago, although they have used the knowledge and skills to some extent ever since. This suggests that these skills and knowledge have a degree of stability over time.

The final reviews for both workbooks can be said to have content validity since they were developed directly from instructional objectives. More important, however, is the question of whether or not the criteria for the achievement of competency in the skills are valid. That is, does the

achievement of competency in these skills say anything about the participants' ability to apply the skills in a practical setting? One way of determining this would be to set up a system whereby the field test participants could be observed carrying out evaluations in the field. Their behaviors relating to "Defining evaluation" and "Dealing with a lack of decision maker cooperation," could be noted and then compared to their scores on the final review. A reasonable match would indicate the validity of the criteria, while discrepancies would suggest that they require modifications.

The design of a complete skill-based curriculum for the training of F-H evaluators is now possible through the use of the final skills list which was discussed in Chapter III. It would have to be decided how many courses would be needed to teach the most important identified skills and whether some can be made optional for those who don't have as deep an interest in evaluation. The planning and development of materials for each course could proceed in the same way that the two workbooks were designed, with more attention being paid to the interaction and overlapping of materials across skills. The field test-revision-field test cycle should be continued until it is decided that there are no more important problems to be found through this process and that a sufficient number of users is achieving competency on each skill the first time through.

The process of developing a complete curriculum is difficult and time-consuming. This helps insure, however, that one can tell what the results of the final review mean, e.g., which objectives were learned and where the instruction and/or the test question is faulty. Because any revisions in the learning materials will eventually be field tested, a variety of changes can be tried out. These might include a change in the mode of instruction, the number of objectives included in one learning package, or the kinds of objectives which are used to operationally define a skill or skills.

The development of curriculum should be continued in this way, both for the benefit of future F-H evaluators and for others, such as classroom teachers, who need an effective way of evaluating instruction on a daily basis. In fact, teachers could profit from learning this process of curriculum development, as much as from the content of this skill-based curriculum. Although time-consuming, this approach to the design of instruction allows for the specific evaluation of the degree of competency achieved. In the long run, it will simplify the teacher's role and improve achievement for students whose problems may not be recognized with a more global approach to teaching and learning.

With respect to the two workbooks developed in the course of this work, the suggested changes should be incorporated into them and the related final reviews. The materials should be field tested again, with two administrations of the

final reviews, as advised above. These field tests should also investigate the value of cycling through the workbook a second time. Thus, a comparison could be made of the question competency achieved on the final review and the degree of improvement when participants who had a chance to review workbook material pertaining to questions for which they had initially failed to earn competency responded to those questions again. The process of field testing should continue until no more major problems can be identified and a sufficiently large number, e.g., 40%, achieve competency on the first try. The materials should then be distributed on a much larger scale for the purpose of training future evaluators and identifying the more minor problems that might have been missed in earlier field tests.

REFERENCES

- Benedict, L. G. (Ed.). The Fortune-Hutchinson evaluation methodology, version I, draft I. Center for Educational Research, University of Massachusetts, Amherst, mimeo, 1973 (a).
- Benedict, L. G. The goals gap in educational evaluation: Identification and development of methodology. Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1973 (b).
- Benedict, L. G. Practical guide for evaluation. Windsor, Conn.: Capitol Region Education Council, 1973 (c).
- Brooks, J. P. Student affairs planning and evaluation handbook. Student Affairs Research and Evaluation Office, Whitmore Bldg., University of Massachusetts, Amherst, mimeo, 1974.
- Center for the Study of Education. Workshop on teacher appraisal for improvement. UCLA Graduate School of Education, Los Angeles, California, 1972.
- Coffing, R. T., Hutchinson, T. E., Thomann, J. B., & Allan, R. G. Self-instructional module for learning the Hutchinson method of operationalizing a goal or intent. University of Massachusetts, Amherst, mimeo, 1971.
- Cronbach, L. J. Evaluation for course improvement. Teacher College Record, 1963, 64, 231-248.
- deCecco, J. F. The psychology of learning and instruction: Educational psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968.
- Far West Laboratory for Educational Research and Development. A training program in educational development, dissemination, and evaluation. Under contract with the National Institute of Education, United States Department of Health, Education, and Welfare, Washington, D. C., March, 1973.
- Gagné, R. N. Learning and instructional sequence. In F. N. Kerlinger (Ed.), Review of research in education. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1973.
- Glass, G. V. The growth of evaluation methodology. University of Colorado, mimeo, March, 1969.

- Goodlad, J. I., Von Stephasius, R., and Klein, M. F. The changing school curriculum. New York: The Fund for the Advancement of Education, 1966.
- Gordon, G. M. A field test of the Fortune-Hutchinson evaluation methodology as it could be employed in the evaluation of national urban league street academies. Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1973.
- Gorth, W. P., O'Reilly, R. P., & Pinsky, P. D. Criterion-referenced evaluation. University of Massachusetts, Amherst, mimeo, June, 1973.
- Guba, E. G., & Gephart, W. J. Training materials for research, development and diffusion training programs. Washington, D. C.: National Center for Educational Research and Development. Final report for grant OEG-1-7-071018-4501, 1970.
- Guba, E. G. & Stufflebeam, D. L. Evaluation: The process of stimulating, aiding, and abetting insightful action. Paper presented at the 2nd Phi Delta Kappa National Symposium for Professors of Educational Research, Boulder, Colorado, November, 1968.
- Hutchinson, T. E. A critique of federal evaluative methodology: A review of what was done compared to what should be done. A paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972 (a).
- Hutchinson, T. E. Class lecture notes, University of Massachusetts, Amherst, 1972 (b), 1973.
- Hutchinson, T. E. Some overlooked implications of the purpose: To provide data for decision making. A paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972 (c).
- Hutchinson, T. E. & Benedict, L. G. The operationalization of fuzzy concepts. University of Massachusetts, Amherst, mimeo, 1970 (a)
- Hutchinson, T. E. & Benedict, L. G. Notes from evaluation workshops. University of Massachusetts, Amherst, 1970 (b).

- Jeffers, E. The interface of measurement and evaluation: Methodological development and field test in an innovative environment. Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1974.
- Kirst, M. W. and Walker, D. F. An analysis of curriculum policy-making. Review of Educational Research, 1971, 41, 479-509.
- Mager, R. F. Preparing instructional objectives. Palo Alto: Fearon Publishers, 1962.
- Merrill, M. D. & Boutwell, R. C. Instructional development: Methodology and research. In F. N. Kerlinger (Ed.), Review of research in education. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1973.
- Mitchell, V. P. Fortune-Hutchinson and me. University of Massachusetts, Amherst, mimeo, December, 1973.
- Percy, V. R. An instructional system for teaching the Fortune-Hutchinson evaluation methodology. University of Massachusetts, Amherst, mimeo, 1973.
- Popham, W. J. & Baker, E. L. Establishing instructional goals. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.
- Rosen, D. J. New evaluation for new schools. Changing schools, An occasional newsletter on alternative public schools. Special issue on evaluation for alternative schools, Bloomington, Indiana: Educational Alternatives Project, Indiana University, May, 1973.
- Rosen, D. J. The Shanti evaluation: A study of the Fortune/Hutchinson evaluation methodology in a public alternative school. Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1974.
- Rosen, D. J., Benedict, L. G. & Hutchinson, T. E. A self-instructional module in the goals process. University of Massachusetts, Amherst, mimeo, 1973.
- Schwab, J. J. The practical: A language for curriculum. School Review, 1969, 78, 1-23.
- Scriven, M. S. The methodology of evaluation. In R. W. Tyler (Ed.), Perspectives of curriculum evaluation. AERA Curriculum Monograph Series, #1, Chicago: Rand McNally, 1967.

- Skinner, B. F. The technology of teaching. New York: Appleton-Century-Crofts Educational Division of the Meredith Corp., 1968.
- Stake, R. E. & Denny, T. Needed concepts and techniques for utilizing more fully the potential of evaluation. In R. W. Tyler (Ed.), Educational evaluation: New roles, new means. The 68th Yearbook of the National Society for the Study of Education, Part II, Chicago: University of Chicago Press, 1969.
- Stetz, F. P. Implementing a "G.A.P." in the Fortune/Hutchinson evaluation methodology. Center for Educational Research, University of Massachusetts, Amherst, mimeo, 1972.
- Thomann, J. B. Workbook for goals-parts integration step of the Fortune-Hutchinson evaluation methodology. University of Massachusetts, Amherst, mimeo, 1972.
- Thomann, J. B. Metamethodology: The first field test. Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1973.
- Tyler, R. W. Basic principles of curriculum and instruction. Chicago: University of Chicago Press, 1950.
- VIMCET Associates. Slide-tape series on evaluation, Los Angeles, California, 1972.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw Hill Book Company, 1962.
- Worthen, B. R., Anderson, R. D., & Byers, M. L. A study of selected factors related to the training of researchers, developers, diffusers, and evaluators in education. Washington, D. C.: American Educational Research Association. Final report for Grant #OEG-0-71-0617 (520), 1971.
- Worthen, B. R. & Sanders, J. R. Educational evaluation: Theory and practice. Worthington, Ohio: Charles A. Jones Publishing Co., 1973.

APPENDICES

APPENDIX A

Evaluation Handbook #1

Defining Evaluation*

Final Review

Facilitator's Handbook to

Defining Evaluation

*Note: The page numbers throughout this handbook have been left as they were in the original field test to provide for ease of reference, and so that the handbook may be used in the future apart from this dissertation.

Evaluation Handbook #1

Defining Evaluation

Introduction:

This booklet contains materials designed to help those interested in evaluating educational programs using the Fortune-Hutchinson Evaluation Methodology (F-H) to understand and be able to explain the definition of evaluation used in F-H.

Evaluators in the field of education often find themselves working with students, teachers, principals and/or superintendents who are relatively unfamiliar with this type of evaluation and either don't know what to expect or may expect something very different from what will be done. Because this evaluation methodology attempts to maximize the use of data collected, it is important that those working with the evaluator fully understand the definition of evaluation, the assumptions behind it, and its implications in terms of the kinds of activities in which the evaluator will and will not be engaged. If there is a disagreement on the part of school personnel to this definition of evaluation, the evaluator should know it as soon as possible and 1) convince them that F-H is effective and should be used; 2) consider using another methodology; or 3) decide not to take the evaluation job.

Instructions:

Read through the booklet and complete the review questions as they occur. If you have any questions, the facilitator will try to help you. You will have an hour. You will be allowed a short break, if you desire, after which the facilitator will hand out the final review. This must be completed without referring to the booklet and will be corrected and handed back by the facilitator. You will have up to one hour for the final review.

The Evaluator Must Be Able to Define Evaluation.

Before setting out to learn how to be an evaluator, one must think about what evaluation means. The definition of evaluation used in the Fortune-Hutchinson Evaluation Methodology is "the providing of data about an enterprise for decision making." In this case we are referring to decision making about an educational program or enterprise of some kind. This definition assumes the following (Benedict, 1973):

- 1) The project personnel have the moral and ethical right to make their own decisions about their own enterprise.
- 2) It is their responsibility, and not that of an outside "expert" or "consultant," to make these decisions.
- 3) The only legitimate purpose of educational evaluation is to provide information to persons making decisions about a project or enterprise to use for whatever means they wish.

These are the major assumptions. Others include: decisions can be made more effectively with appropriate data; data should be used by decision makers in their decision making processes; the goals the decision makers have for the project are the best guides as to which data are most important to them and most likely to be used; and evaluation is an ongoing process which should be built into the program from the beginning, if possible, in order that informed decisions may be continually made on the basis of evaluation data.

Review #1

Which of the following is the correct F-H definition of evaluation? Check one.

1. To determine the worth of the program.
2. To provide data about an enterprise for decision making.
3. To compare the program with others like it.
4. To look objectively at what the program is doing.

If you chose #2, you are correct. Congratulations! Go on to page 5. If you answered incorrectly, please reread page 2 before going on to page 5.

Review #2

Give four assumptions behind this definition of evaluation. Write them below in the space provided.

Here is the complete list of assumptions:

1. The project personnel have the moral and ethical right to make their own decisions about their own enterprise.
2. It is their responsibility, and not that of an outside "expert" or "consultant" to make these decisions.
3. The only legitimate purpose of educational evaluation is to provide information to persons making decisions about a project or enterprise to use for whatever means they wish.
4. Decisions can be made more effectively with appropriate data.
5. Data should be used by decision makers in their decision making processes.
6. The goals the decision makers have for the program are the best guides to which data are most important to them and most likely to be used.
7. Evaluation is an ongoing process which should be built into the program from the beginning, if possible, in order that informed decisions may be continually made on the basis of evaluation data.

If you did not get four answers correct, reread page 2 before going on to page eight.

Other Approaches to Evaluation:

The following are other reasons frequently given for doing evaluation (Benedict, 1973):

- 1) For public relations reasons--to improve the program's image, to gain funding, etc.
- 2) To find out what some group needs.
- 3) To make program or planning decisions.

These are not reasons for using F-H, however. The skills needed to do a good public relations campaign are not necessarily those needed by an F-H evaluator. The purposes of these two endeavors are different, although information gained by doing an evaluation may be used for public relations reasons.* If one wishes to become adept in the public relations field, he or she should study public relations skills, rather than F-H evaluation skills.

The second reason given above for doing an evaluation is also not accomplished by F-H, as it is a needs analysis, rather than an evaluation, which will give one information about needs. As with the public relations example, the purposes involved and skills required to perform a needs analysis and an evaluation are different, and one should not study the latter in order to do the former. An evaluator should make sure that the program decision makers don't expect a needs analysis, instead of an evaluation.

The third reason listed above, "to make program or planning decisions" is also not the purpose of F-H. The making of decisions is a job solely for decision makers. If the quality of decisions made needs improvement, the decision makers should try

* Evaluation data should not, however, be screened for its "acceptability."

to obtain better decision making skills, rather than F-H evaluation skills. Very often an evaluator is not trained to help improve decision making skills, and, as with the needs analysis example above, the evaluator should be sure beforehand just what it is the decision makers expect. They, in turn, should understand what they will and will not be getting.

Review #3

Answer the following question in the space below.

Why should F-H evaluation not be equated with public relations, needs analysis, or decision making?

If you said that these four processes should not be equated because they all have different purposes and require different skills, that's correct. Very good! Please go on to the next page. If you answered the question incorrectly and do not understand the answer, please review page 7 before going to to page 11.

Not only is the Fortune-Hutchinson Evaluation Methodology not designed to do the three things mentioned above, but it is also a break with traditional models of evaluation. Evaluation has frequently been thought of with a great deal of anxiety in the past, especially by directors of educational programs and enterprises, because it has traditionally meant that a person or persons outside the program would pay an "on-site" visit to the program location and, in a comparatively short period of time, would place a value on the program which could profoundly affect its future. The values of the "evaluator-as-expert" or "board-of-experts" were what determined how the program would be viewed and what was measured. Whether or not those values were shared by the decision makers in the program would determine whether or not the data collected were useful to the program staff. If the information gathered during this type of evaluation had any relevance to the program decision makers' goals, it was often by chance. An example of this type of assessment is the accreditation process for which a team brings with them criteria which deal mostly with the physical resources of the educational setting, rather than with the more important goals that educators have.

Another model, the "outside evaluator" model has been used extensively by the Office of Education in Washington, D. C. An individual with a background in education measurement, statistics, and research design approaches the evaluation problem with standardized tests and sets up a quasi-control group to study. (Because the program is probably already under way when the evaluator is called in, it is impossible to set up a true control group which would have required the random assignment of

individuals to the treatment and control groups.) Generally, a pre- and post-test are administered to the two groups, frequently no significant differences are found, and the data collected are written up in a final report which is shelved and never looked at again.

Other evaluation models include Stufflebeam's "CIPP" model and Provus' "Discrepancy" model. The former stands for "context, input, process, and product" and is concerned with relating educational outcomes to their antecedent processes as well as to the objectives for the program. The latter attempts to describe the discrepancies between the outcomes of a program and what "should be," given the program goals. Both models have grown from the definition of evaluation put forth by Lee J. Cronbach in 1963, i.e. "...the collection and use of information to make decisions about an educational program," and can be said to be included in the broader "decision maker" model of evaluation, along with the Fortune-Hutchinson Evaluation Methodology. Models, however, have a major disadvantage over methodologies in that often their recommended procedures are not stated in operational terms so that the evaluator has to guess about how to accomplish them (Hutchinson, 1972).

Review #4

Please match the model with its correct characteristics. Write your answers below.

1. accreditation team
 2. outside expert
 3. CIPP model
 4. discrepancy model
- a. often uses standardized tests and quasi-control groups
 - b. determines where program goals and program outcomes fail to coincide
 - c. brings along own evaluation criteria which deal mostly with physical resources
 - d. relates educational outcomes to antecedent processes and to program objectives

The correct answers to Review #4 are: 1-c; 2-a; 3-d; and 4-b. If you got them all right, that's very good. Please go on to page 15. If not, please go over pages 11-12 again until you understand the answer. Then proceed to page 15.

What F-H Isn't:

We have briefly discussed the definition of F-H and some of the assumptions behind it. We have also cited a few of the existing models of evaluation in contrast. Without going into a thorough discussion of F-H at this point, there are two more comparisons we should make.

First, evaluation is not research. The purpose of providing data for decision making is not the same as the purpose of generating knowledge. Therefore, the means of achieving these two purposes will differ. An evaluation requires continual data to be obtained during the course of the program, rather than post hoc data which is the case in an experimental situation (Guba and Stufflebeam, 1968). The concern of F-H is in providing specific information to decision makers so that they can determine to what degree their goals are being accomplished, while that of classical research design is to generalize results. Other factors such as: holding the treatment intact during the experiment; limiting attention to a few, easily controlled variables; and the practice of random assignment are in conflict with evaluation in general and F-H in particular. Evaluation assumes that the treatment (the educational program) will change for the better because of the feed-back of evaluation data, and it is accepted that the many interacting variables can't be controlled. Also, random assignment is usually not possible, as the evaluator is often called in after the program has been operating for a while, or ethical considerations do not permit the withholding of a special program from a designated "control" group (Guba and Stufflebeam, 1968).

Secondly, evaluation isn't the same thing as measurement, although measurement is a part of evaluation. Most often in F-H, observational techniques are developed from components of the goal statements of decision makers, although standardized tests may be used if they pertain especially well to the goal-component in question. The decision maker is directly involved with all processes leading up to the design of the observational techniques, and especially in the process of breaking down his or her goals into operational components. In this sense, F-H may be considered "subjective" because all processes rely heavily on each individual decision maker's goals and specification of those goals.

Review #5

In the blanks below, put the number of the statements which apply to research and to evaluation, respectively.

research _____

evaluation _____

1. specific information on goal accomplishment
2. treatment changes
3. random assignment
4. many interacting variables
5. generalizing results
6. post hoc data
7. holding treatment intact
8. easily controlled number of variables
9. continual data

The answers are:

research 3, 5, 6, 7, 8

evaluation 1, 2, 4, 9

If you had trouble with any of the above, please review pages 15-16 before going on.

Review #6

From what are the observational techniques in F-H developed?

Answer below.

The answer to Review #6 is that observational techniques in F-H are developed from the components of the goal statements of decision makers. If you answered this correctly, that's very good. Please go on to page 21. Otherwise, please review the last paragraph on page 16 before going on.

Implication of the Purpose:

The most important implication of the purpose "to provide data about an enterprise for decision making" is that information provided by the evaluation should be used. It should be used to make decisions about the program by those with whom the evaluator has been working throughout the evaluation. Thus, if Fortune-Hutchinson is to be a successful evaluation methodology, it must produce evaluations which furnish information which is used by those decision makers who wanted it in the beginning.

This implication of the purpose has led to the identification of three criteria to determine the degree of success of a Fortune-Hutchinson evaluation (Hutchinson, 1972). The criterion of efficiency is used to determine what proportion of the information generated by the evaluation was used by a decision maker to make a decision about the program. The more information gathered which was not so used, the less efficient the evaluation was.

Secondly, the criterion of completeness is used to determine the proportion of decisions made by decision makers in the program which were made with the use of evaluation-generated information. The greater the number of decisions made without evaluation data, the less complete the evaluation was. Usually an evaluation can only be 100% complete if there is a very large budget available and the evaluator can do a really thorough job. Completeness should be as high as possible, however, considering resources available.

Finally, the criterion of focus indicates that, when the evaluation is not going to be 100% complete, the evaluator should try to provide the decision maker(s) with information for his or her most important, as opposed to least, important decisions. This is saying

again that resources for the evaluation should be used in the best possible way: to provide as much usable data as possible; to provide data for as many decisions as possible; and to provide data for important, rather than unimportant, decisions when such a choice must be made.

Review #7

Match each criterion with its appropriate description. Answer in the space below.

1. completeness
 2. efficiency
 3. focus
- a. proportion of information generated by evaluation used to make decisions
 - b. proportion of information provided by the evaluation for the most, rather than the least, important decisions the decision maker makes
 - c. proportion of decisions made with evaluation-generated information

The correct answers to the matching question on page 23 are:

1-c; 2-a; and 3-b.

If you failed to match these questions correctly, please review page 21 and page 22. Otherwise, please continue with page 25.

There are processes built into the Fortune-Hutchinson methodology all along which promote the achievement of the three criteria discussed above.* The goals and parts processes, tests of completeness, and the process of putting lists into priority order, help to insure that the information gathered during the evaluation will have decision maker validity, i.e. that the decision maker will consider it valid, important, and, therefore, usable. The goals process elicits the goals of the individual decision makers, and the parts process helps him or her identify the working parts of the program. When the two lists are cross-matched, each decision maker's identified goal related to the part(s) of the program which should be accomplishing that goal, and each part of the program which each decision maker has identified relates to the goal(s) which it should be accomplishing. This indicates very specifically what information is going to be most important to this decision maker and helps to promote an efficient evaluation.

In addition to the goals and parts processes, the methodology encourages the use of so-called "tests of completeness." This term means that, throughout the evaluation, each decision maker is provided with points of view other than his or her own (e.g. other decision makers' views or those of other school personnel, experts in the field, etc.) to react to in order that the decision maker's thinking may be further stimulated. This helps the decision maker be as complete as possible in the identification of goals for the

*The major steps of F-H are: goals process; parts process; matching of goals and parts; operationalization of goals; design of observational techniques; implementation of observational techniques; reporting of data; evaluation of evaluation; and redesign.

program, parts of the program, and in all other aspects of the evaluation. This, in turn, helps insure that important information that the decision maker needs won't be omitted, and that the evaluation will be as complete as possible.

One other characteristic of the Fortune-Hutchinson methodology is important to mention here, and that is the process of putting lists into priority order. The evaluator suggests to the decision maker several criteria by which a list may be put in order, and the decision maker may choose one or more of these or offer one of his or her own. It is the decision maker, however, who actually puts the lists in order. Here again this helps insure that the information collected will be important to the decision maker, rather than trivial, in the case where resources are limited and not all desirable information can be collected within given time and money constraints. This helps the evaluation rate high on the criterion of focus.

Review #8

The goals and parts processes, tests of completeness, and the process of putting lists into priority order together help to insure which one of the following? Check one.

- a) objectivity
- b) reliability
- c) continuity
- d) decision maker validity

The correct answer to the previous question was "d", decision maker validity. If you answered it correctly, that's very good. Please go on to the final question on page 29. If you did not get the right answer, just remember that "decision maker validity" means that the decision maker perceives the data as being appropriate to his or her goals and data needs. If decision maker validity is maximized, this will help insure that evaluation-produced data will actually be used for decision making purposes. Please go on to page 29.

Review #9

Match each criterion with the process designed to accomplish it.

Answer in the space below.

- | | |
|-----------------|------------------------------|
| 1. efficiency | a. test of completeness |
| 2. completeness | b. priority order |
| 3. focus | c. goals and parts processes |

The correct answer to the previous matching question was: 1-c; 2-a; 3-b. If you answered this correctly, congratulations! You have successfully completed the content material and the review material. You are now ready to complete the final review test which follows. If you did not answer this question correctly, please go back over pp. 25-26 until you feel sure that you understand your error. Then proceed to the final review test.

References

- Benedict, L. G. Practical guide for evaluation. Windsor, Conn.: Capitol Region Education Council, 1973.
- Cronbach, L. J. Evaluation for course improvement. Teachers College Record, 1963, 64, 231-248.
- Guba, E. G. & Stufflebeam, D. L. Evaluation: The process of stimulating, aiding, and abetting insightful action. Paper presented at the 2nd Phi Delta Kappa National Symposium for Professors of Educational Research, Boulder, Colorado, November, 1968.
- Hutchinson, T. E. Some overlooked implications of the purpose to provide data for decision making. Paper presented at the annual meeting of the American Research Association, Chicago, 1972.
- Hutchinson, T. E. Class lecture notes, 1972, 1973.

Final Review

1. Give the F-H definition of evaluation. Answer on separate paper.
2. Check any of the following which are assumptions behind the F-H definition of evaluation.
 - a. An evaluation can best be carried out without decision maker interference.
 - b. Evaluation is an ongoing process which should be built into the program from the beginning, if possible, in order that informed decisions may be continually made on the basis of evaluation data.
 - c. Evaluation should improve the program's image.
 - d. Evaluation should be free of the constraints of the decision makers' goals.
 - e. The project personnel have the moral and ethical right to make their own decisions about their own enterprise.
 - f. An evaluation should make as much use as possible of classical research design.
 - g. The only legitimate purpose of educational evaluation is to provide information to persons making decisions about the project to use for whatever means they wish.
 - h. It is the responsibility of project personnel, and not that of an outside "expert" or "consultant" to make the decisions affecting their program.
 - i. Evaluation should only make use of standardized tests.
 - j. An evaluation should be an assessment of a program in terms of an objective standard of worth.

k. The goals the decision makers have for the program are the best guides to which data are most important to them and most likely to be used.

l. Decisions can be made more effectively with appropriate data.

m. Decision makers should make decisions on the basis of the evaluator's opinions about the program.

n. Data should be used by decision makers in their decision making processes.

3. As evaluator, give a brief but specific answer (in terms of F-H) to each of the following decision maker questions or statements. Answer on separate paper.

a. "What should we do about our program?"

b. "We don't want to collect that kind of data--it would be embarrassing."

c. "Our main goal is to identify the students' needs."

d. "We want to verify the worth of our program."

e. "How can we achieve our goals?"

f. "Can't you just tell us what goals we should have?"

g. "We want to make better decisions about our program."

4. For the next two questions, select the one major problem with each model.

a) evaluator-as-expert model

1 - involves too much of the decision makers' time

2 - it is difficult to specify the program "content"

3 - evaluation reports are over-simplified

4 - use of standard evaluation criteria which might not really apply to a given program

b) outside evaluator model

1 - use of standardized tests which might be inappropriate for the particular program

2 - development of specific data-collection instruments is time consuming

3 - outside evaluator may become personally involved in the program

4 - provides only public-relations type of data

5. What is the major problem of models, as opposed to methodologies, in general? Answer on separate paper.

6. What do the CIPP and Discrepancy models of evaluation have in common with F-H? Answer on separate paper.

7a. Name the three criteria developed to evaluate an F-H evaluation.

b. Define each of the above. Answer on separate paper

8. The next three questions concern the importance of the following processes to F-H. Check as many as apply for each.

a) goals and parts processes

1 - promote the criterion of efficiency

2 - promote the criterion of focus

3 - provide specific information of interest to the decision maker

4 - let decision maker see what the experts would do with the program

b) tests of completeness

1 - promote the criterion of focus

2 - can help stimulate decision maker's thinking

3 - promote the criterion of completeness

4 - do away with the need for standardized tests

c) process of putting lists in priority order

1 - promotes the criterion of completeness

2 - promotes the criterion of focus

3 - helps prevent decision makers from using each other's
lists and, therefore, data

4 - helps conserve resources

Facilitator's Handbook to
Defining Evaluation

Instructions to Facilitator

Allow up to an hour for the completion of this booklet. Tell students that if they have finished, they may review the materials until it is time for the final review test and that they will not be allowed to look back at the booklet while the final review is in progress. As soon as a student wishes to take the final review, he or she may signal this to the facilitator who will then hand out the tests along with some paper, and ask that booklets be closed. Everyone should have started the test after an hour. If a break is desired at this time, have everyone start together after the break at the beginning of the second hour.

Once the test is completed, students should turn them back to the facilitator who should score them according to the directions that follow and hand them back to students either on the same day or soon thereafter. Students who have failed to achieve competency should be informed to review those sections of the booklet listed at the end of this manual relating to the appropriate questions. They may then attempt the final review again when they feel ready. Other students who have achieved competency should also review any sections relating to questions they missed, although it isn't necessary for them to take the test again.

Another way of using these materials is to have the students go through the booklet at home for homework and to take the final review together in class the next day.

Scoring the Final Review

Question #:

1. The only acceptable answers are: "to provide data about an enterprise for decision making" or "to provide data for decision making."

Another word for "provide" such as "furnish" or "collect" may be used. Also, "program," "project," etc. may be used for "enterprise."

The correct answer is worth 5 points. Otherwise, no credit will be given.

2. The correct answers are: b, e, g, h, k, l, and n.

One point should be given for each correct answer and -1 for each incorrect answer chosen.

- 3a. The answer should include a brief indication of the roles of the evaluator and the decision maker. Allow one point for each included.

Examples

Full credit (2 points) for the following:

- 1) It is the decision maker's job to make decisions about the program. The evaluator merely helps provide the best data for facilitating this.
- 2) The evaluator designs and carries out the evaluation which provides the information to meet the needs of the decision maker. Only the decision maker can make decisions about the program.

Half credit (1 point) for the following:

- 1) Decision makers must make their own decisions.

Question #:

- 2) Evaluators don't make decisions about the program.
- 3) Evaluators provide data for decision making.

No credit for the following:

- 1) Do what you think is best.
- 2) I can't give you that information.
- 3) Tell me about your program and I'll try to tell you what to do about it.

3b. The answer should say that the collection of positive data only is a job for a public relations expert. Also, the answer should have one other point--either that negative data is helpful as a basis for program improvement or that ignoring possible areas of malfunction in the program may cause the program to fail.

Full credit (2 points) for the following:

- 1) The evaluator hasn't been hired to provide only favorable data to decision makers. This is a job for a public relations expert. Decision makers should be able to make good use of negative data to improve their program

Half credit (1 point) for the following:

- 1) Bad data is not necessarily an indication of a poor program. Decision makers should realize this and try to improve their program.
- 2) The evaluator is not a public relations expert.

No credit for the following:

- 1) How do you know that data would be embarrassing?
- 2) O.K. We'll concentrate on other areas of the program.

3c. The answer should say that this may indicate that the decision makers want a needs analysis. Since the evaluator may

Question #:

not be equipped to do this for them, it should be determined exactly what job the decision makers want done.

Full credit (2 points) for the following:

1) This is a noteworthy goal, but the evaluator's job does not usually include doing a needs analysis as well as an evaluation. The extent of the evaluation job should be determined ahead of time, and, if the decision makers really want a needs analysis, the evaluator should turn down the job unless he/she has the ability and desire to do what is asked.

No credit for:

- 1) That is a fine goal! What are you doing about it?
- 2) We must make up a questionnaire and give it to students to determine their needs.

3d. This question, like 3b, is getting at the real desire for public relations information rather than a free open evaluation.

Full credit (2 points):

- 1) The decision makers here are interested in positive data only. They want a public relations effort, rather than an evaluation.
- 2) You have to be sure that your program has "worth" before you go about verifying it. If you aren't interested in first determining the real worth of your program, you are looking for public relations, not evaluation.

No credit:

- 1) What is the worth of your program?
- 2) We must first identify your goals.

Question #:

3e. The point here is that the decision makers are the best judges of this.

Full credit (2 points):

- 1) It isn't up to the evaluator to determine how goals may best be achieved, but only to provide the data on the accomplishment of the goals. If the decision makers aren't satisfied with the level of accomplishment, it is up to them to decide how to improve it.
- 2) This may only be determined by the decision makers. The evaluator doesn't have the right to make these decisions.

No credit:

- 1) We must first find out to what extent you are achieving them.
- 2) If you don't know, I can't tell you.

3f. Again, this is the role of the decision makers, not the evaluator. In addition, the answer might include that the decision makers' goals are the best guide to what data they want and will use.

Full credit (2 points) for:

- 1) The evaluator shouldn't do this because F-H wants to make sure that decision makers will use the data generated by the evaluation. Only the decision makers' own goals can help insure that data (relating to these goals) will be used.
- 2) This isn't the responsibility of the evaluator and not in the best interests of the decision makers or the program. The evaluator's goals may be irrelevant to the decision.

Question #:

makers. They could then righteously ignore any data coming out of such an evaluation.

Half credit (1 point):

- 1) This is the decision makers', not the evaluator's, role in F-H.
- 2) The evaluator's goals are not the best guide to how the program should operate.

No credit:

- 1) You should know your goals already.
- 2) How can you run a program if you don't have any goals?

3g. The answer should state that it isn't the job of F-H to improve decision making skills among project personnel. This should be clarified because, unless the evaluator wishes to work in this area, he/she should refuse to do so.

Full credit (2 points) for:

- 1) This is a worthwhile desire, but all I can do is to provide you with the data you want most. You must do the actual decision making.
- 2) I can advise you of some alternatives, but you must make the decisions yourself. You will have to look elsewhere for someone to actually assist you in improving your decision making skills. I don't have the time to do that and the evaluation, too.

No credit:

- 1) I hope this evaluation will help you do that.
- 2) What data do you think can best help you do this?

4a. 4 (1 credit if right, none if wrong)

Question #:

4b. 1 (1 credit if right, none if wrong)

5. Models tend to be descriptive rather than prescriptive.

or

Their recommended procedures aren't stated in operational terms so that the evaluator has to guess about how to accomplish them.

The answer must have this basic idea, or no credit will be given. If correct, give one point.

6. All are included in the "decision maker" model of evaluation.

or

All stress the role of the decision maker.

The answer must have this basic idea. One credit if right, none if wrong.

7a. Focus, completeness, efficiency.

One point given for each criterion named correctly.

No credit for wrong answers.

7b. Correct definitions (give 1 credit each--none for wrong answers):

efficiency: the proportion of data provided by the evaluation which are used to make decisions.

completeness: the proportion of decisions made with evaluation-generated data.

focus: the proportion of important vs. unimportant decisions made with evaluation-generated data.

No credit for:

efficiency: all information provided by the evaluation is used to make decisions about the program

Question #:

or

all the information the decision maker wants
is used

completeness: all decisions are made with evaluation gener-
ated data

or

all decisions take the evaluation into con-
sideration

focus: all important, as opposed to unimportant deci-
sions are made using evaluation generated data

or

important decisions are made with the evalua-
tion in mind

Note: "information" may be used for "data" and "provided" or
"provided by the evaluation" may be used for "generated"
or "evaluation-generated."

8a. 1 and 3

8b. 2 and 3

8c. 2 and 4

Under #8, one point is given for each correct answer and -1
for each incorrect answer.

Determining Competency in this Skill:

In order for someone to earn a "pass" on this skill, he or she must have the following correct:

Question #1	5	points	out	of	a	possible	5
#2	5	"	"	"	"	"	7
#3	10	"	"	"	"	"	14
#6	1	"	"	"	"	"	1
#7	5	"	"	"	"	"	6
#8	3	"	"	"	"	"	6

Note: for question #8, at least one point must be earned in each section (8a, 8b, and 8c). Thus, if one section yields a "-1", the points may not be made up by a "2" in each of the other two sections. Questions #4 and #5 are not considered necessary for competency on this skill. They are worth 2 points and 1 point. The totals will be, therefore, 29 or better out of a respectively possible 42, with the above distribution.

Remedial Materials:

Inform students who have missed questions to review the appropriate pages in the booklet listed below.

<u>Question #</u>	<u>Page or Pages in Booklet</u>
1	2
2	2, 5-16
3a	2
3b	7
3c	7
3d	7
3e	2
3f	2
3g	7-8
4a	11
4b	11-12

<u>Question #</u>	<u>Page or Pages in Booklet</u>
5	12
6	12
7a	21
7b	21-22
8a	25
8b	25-26
8c	26

If a student earns less than 29 points overall, he or she should review the whole booklet and take the final review again. If a student only misses one or more of questions 1-3 or 6-8 or earns less than the required number of points on any of these questions, he or she may review just the questions missed and be re-tested on just those questions.

APPENDIX B

Evaluation Handbook #2*

Dealing with a Lack of Decision Maker Cooperation

Final Review

Facilitator's Handbook to

Dealing with a Lack of Decision Maker Cooperation

*Note: The page numbers throughout this handbook have been left as they were in the original field test to provide for ease of reference, and so that the handbook may be used in the future apart from this dissertation.

Evaluation Handbook #2

Dealing with a Lack of Decision Maker Cooperation

Introduction

While many "un-cooperative" decision maker behaviors can have more than one cause, the evaluator experiencing the lack of cooperation must try to diagnose the problem(s) and find solutions. Each situation described below may indicate the presence of more than one problem. The important thing, however, is for the evaluator to recognize a lack of cooperation when it occurs and to try to alleviate it. It is also important for the evaluator to know that the problem(s) may not disappear, despite his/her efforts, and, in that case, to consider the following: discussing the problem with the decision maker in question, specifically requesting more cooperation or a more manifest commitment to the evaluation; reporting the lack of cooperation to the contract decision maker and asking that the decision maker be replaced with someone more cooperative or that the contract decision maker take some other action to improve the situation; or, if these attempts fail, giving up the evaluation effort. It would be a good idea for the evaluator who suspects that sufficient cooperation is being withheld, to keep a log documenting these behaviors so that, should a confrontation become necessary, there would be a record for the evaluator to fall back on.

General Form of Workbook

Each "problem" will be treated in five sections as follows:

- A Scenario in which a lack of decision maker cooperation is occurring.
- B Probable problem diagnosis.
- C Ways to anticipate and avoid the problem in the future.
- D Ways to deal with the problem when it occurs.
- E Other possible problems.

Instructions

Please read through the workbook, answering the review questions as they occur. After reading the ten problems included herein and studying the summaries, you will take the final review during which you should not refer to the booklet. Allow an hour for reading the materials and half an hour for the final review.

Review #1

What is one thing an evaluator might try if his/her efforts to alleviate a lack of decision maker cooperation do not seem to be working? Write your answer in the space below.

A possible course of action might be to do one or more of the following:

1. Discuss the problem with the decision maker.
2. Discuss the problem with the contract decision maker.
3. Give up the evaluation effort.

If you failed to include one of the above in your answer, please review page 1. Otherwise, please go on to page 5.

Problem I: A lack of decision maker commitment due to a lack of time.

A As an evaluator you are finding that things are going smoothly with the evaluations of a sixth grade reading program, with the exception of one decision maker, Miss Y, who seems to be impossible to reach. When you leave messages, Miss Y does not return calls, and, when an appointment is made, Miss Y inevitably has to break it or change it. Miss Y often says that she just doesn't have the time for the lengthy procedures required by the Fortune-Hutchinson (F-H) evaluation. When you do meet, Miss Y always tries to short-cut the steps and frequently asks, "Is that all now?" If you leave some paperwork with her to finish for the next meeting, she often has not had a chance to get to it. She seems well-meaning and conscientious, but she cannot seem to devote even a minimal amount of time each week to evaluation activities.

B This problem, that the decision maker does not have enough time to work on evaluation activities, is one which is frequently voiced. The evaluator should beware here, however, because this may be a ploy on the part of the decision maker to cover up a lack of cooperation due to other reasons. If, however, the decision maker seems to agree with the philosophy of the evaluation and to be as conscientious as possible about carrying out tasks, the problem should be seen as a squeeze on the time and energies of the decision maker which, perhaps, was not anticipated by the contract decision maker when designating this person to participate in the evaluation.

C This problem should be anticipated by the evaluator, especially since F-H, more than any other evaluation model or methodology, demands a heavy commitment of time from a decision maker desiring to have meaningful data about a program. The following procedures may be followed or considered.

- 1) Once the contract has been negotiated, the evaluator should speak with each designated decision maker, giving him, her, or them a broad outline of F-H and an idea of what participation in the evaluation will entail. Strong negative reactions from anyone at this point should be talked out, and, if they can't be resolved, the contract decision maker should be made aware of it.
- 2) When the decision maker understands that the evaluation will involve a time commitment, he or she should be prepared to give an estimate of the minimum amount of time which can be guaranteed for the evaluation each week. (This should not include time needed for severe crisis situations which cannot be foreseen.) The evaluator should make a note of this commitment for his or her planning of activities.
- 3) If the decision maker has, normally, a very busy schedule, it might be a good idea to identify, from the beginning, someone who shares the decision maker's goals for the program as much as possible and can fill in for the decision maker when time is short. This person is known as a "surrogate," and the decision maker's approval must be secured before the surrogate can participate in the evaluation.

D Once the evaluation is underway and the time problem begins to emerge, there are a few other ways an evaluator might handle it.

- 1) The evaluator can act as surrogate to a limited degree, being sure to get decision maker approval any time the evaluator completes a process for the decision maker. This situation should be avoided whenever possible, as some decision makers are afraid to disapprove of something the evaluator has done.
- 2) There is a "short-forms" approach to goals which was developed by Rosen which produces a goals list which may not be as complete as possible but which is produced in a fraction of the normal required time. "Short-forms" procedures for other major steps of F-H should be developed and used when decision maker time is tight.
- 3) The evaluator can point out the problem to the decision maker and try to obtain the originally agreed-upon minimum number of hours per week, or satisfactory compromise.
- 4) Some major steps or substeps of F-H may be eliminated if necessary. Such steps may or may not be so designated in the methodology itself. The evaluator should attempt to fill in gaps where necessary.
- 5) The evaluator might convince the decision maker that more funds are needed for evaluation in order to free more of the decision maker's time for these activities.

E Because a lack of time is the most common reason given by decision makers who don't wish to cooperate with the F-H

procedures for other reasons, the evaluator should realize that the apparent problem in this case may not be the real problem. The evaluator might consider that the real problem may be numbers II, III, IV, V, VI, VIII, or X, or a combination of any of these. A casual talk with the decision maker about what he/she would really like to get out of evaluation or about whether or not the evaluation is meeting the decision maker's expectations might provide the evaluator with clues as to which problem is operating here. Otherwise, the evaluator will just have to try to determine what the symptoms of the lack of decision maker cooperation are, and with which problem they seem to fit best.

Review #2

- A. If a lack of time seems to be the problem disturbing the decision maker, give two ways of alleviating it. Answer in the space below.
- B. What is one way that you, as evaluator, can avoid the problem of a lack of decision maker time? Answer in the space below.

- A. Ways of alleviating this problem include:
1. Evaluator can act as "surrogate" to a limited degree.
 2. Use of the "short forms" approach.
 3. Discussing the problem with the decision maker and trying to obtain the agreed-upon minimum number of hours or a satisfactory compromise.
 4. Eliminate some major steps or sub-steps of F-H.
 5. Convince the decision maker that more funds are needed for evaluation.
- B. Ways of avoiding this problem may include the following:
1. Giving decision makers a good idea, from the beginning, of what participation in F-H will entail.
 2. Obtaining a guaranteed time commitment of a minimum number of hours per week from each decision maker.
 3. Identifying a "surrogate" for a very busy decision maker from the beginning.

If you answered correctly two of the choices under "A" and one under "B," that's very good. Please continue with page 11. If you did not get these right, please review pages 6 and 7 before continuing with the booklet.

Problem II: A lack of commitment due to the decision makers not wanting to know what they're doing wrong.

A You are carrying out an evaluation of Project Chance using F-H. You are puzzled about the behavior of Mrs. S, however, a decision maker, who, although she has worked on the evaluation procedures as long as other decision makers, has accomplished less than half of what they have done. Mrs. S seems to be dragging her feet. She has to know exactly what the results of each activity will be used for and frequently wants to "do over" a list, saying she wasn't aware at the time that it would play such an important role in the evaluation. Mrs. S seems to be trying to play games with F-H, i.e., attempting to identify and specify goals which she is already sure are being achieved. You sometimes get the feeling that she doesn't trust you, especially because Mrs. S doesn't admit to the existence of any problems with the program. She is convinced that everything is going fine and seems to resent your mentioning the problems that other decision makers see. In fact, she doesn't seem to believe that the other decision makers could feel this way and tends to ignore it. She seems perfectly happy to close her eyes to any problem areas that might exist.

B The problem of the decision maker who doesn't want to recognize the problems of the program is also fairly common. This decision maker would prefer either not to have an evaluation at all or to have some kind of public relations effort during which only information favorable to the program would be produced.

C It is often very difficult to discover that a decision maker feels this way until it is too late to do much about it. Two suggestions for anticipating this problem follow.

- 1) When the evaluator initially speaks with each of the decision makers, he/she should ask each of them what he/she feels is the major problem facing the program. Anyone who feels there are no problems might be asked what kinds of data he/she hopes to obtain from the evaluation. If the response indicates that the decision maker is looking for a verification of all the positive qualities that the decision maker assumes the program to have, the evaluator should be aware that this person may not really want an F-H evaluation.
- 2) The evaluator can try to insure that the decision makers feel minimally threatened by an F-H evaluation. The evaluator should stress that even data indicating that certain goals aren't being achieved are very useful because they can help program personnel decide whether alternative strategies should be undertaken in order to achieve the goals or whether the goals are unrealistic and should be dropped.

D When the evaluation is underway, the evaluator can look for other signs that a decision maker doesn't want an F-H Evaluation.

- 1) If the decision maker's goals seem to fall in the realm of "documenting the value" or "verifying the worth" of the program, the evaluator should re-emphasize to the

decision maker that the "value" or "worth" must be identified before it can be documented or verified. If the decision maker still fails to understand this, the best course of action might be for the evaluator to discuss the problem with the contract decision maker.

- 2) If the decision maker's goals on the final list are chosen largely from the test of completeness materials, especially from the goals of high-ranking decision makers in the program, it is possible that the decision maker is trying to gain approval and isn't being completely honest about his or her goals. The evaluator might try immediately to break down one of the goals with the decision maker to see whether he or she really has operational meanings for the goal(s). This will save time in the long run because it will avoid the in-between processes in case the decision maker has been playing games. In that case, the procedure mentioned under "1" above may be used.
- 3) Other indications of this problem may be a decision maker's desire to edit any evaluation reports, especially those which might also go to a higher decision maker or to a funding agency and which contain data he or she feels are uncomplimentary to the program. The decision maker may also want to avoid the collection of certain kinds of data--either from his or her own goals list or from those of other decision makers. The evaluator should beware of false excuses about why some data can't or shouldn't be

collected and, where possible, should check this with another source.

- 4) Finally, if goals lists or other lists or statements are repeatedly changed at the decision maker's request, there is a chance that this problem of not wanting an F-H Evaluation is present. The evaluator should note whether the changes are substantive or whether they seem to be more re-wordings of the originals. If the latter is occurring, the evaluator should consider that the decision maker may be insecure and stalling for time, afraid to proceed through the steps of F-H in an open manner. Again, bringing the problem into the open seems to be the best way to deal with it, and, if this fails, it should be discussed with the contract decision maker.

E Other possibilities here would be: a lack of decision maker cooperation due to a lack of understanding of the methodology (III); the decision maker not being "methodologically oriented," (V); the decision maker having a status hang-up (VI); the decision maker being insecure or under pressure (VIII); or the decision maker having been wrongfully identified from the beginning (X). The evaluator should investigate these possibilities if the initial efforts fail.

Review #3

A. The decision maker seems absolutely convinced beforehand that his or her program is great; refuses to believe that other decision makers have identified problem areas; and resents you, as evaluator, suggesting that the accomplishment of some goals has not been established. The problem existing here is most likely (please fill in the problem in the space below).

B. Give one way of avoiding this problem and two ways to help solve it in the space below.

- A. This problem is most likely that the decision makers in the program don't want to know what they're doing wrong.
- B. Ways of avoiding this problem include:
1. Discovering at the beginning any decision makers who believe there are no problems facing the program and any who want the evaluation to provide only positive data.
 2. The evaluator can try to be minimally threatening to the decision makers and to let them know that negative data can be useful for program improvement.
- C. Ways of alleviating this problem include:
1. Explaining to the decision maker that goals which suggest "documenting the value" or "verifying the worth" of a program assume that the value or worth has already been identified.
 2. Evaluator may discuss the problem with the contract decision maker.
 3. Evaluator might try an immediate breakdown of a decision maker goal if most of these have come from the test of completeness. This will tell the evaluator whether the decision maker really can specify the goal or whether he/she has been trying to play games with F-H.
 4. Evaluator should be aware of false excuses about why some data can't or shouldn't be collected. These reasons should be checked with another source.
 5. If possible, the contract decision maker should be made aware of any decision maker attempt to edit evaluation

reports. In the case of the contract decision maker trying to do this, the evaluator should make it clear that this is not acceptable.

6. Frequent changes of goals lists or other lists should be examined carefully. If they are mere re-wordings of the originals, the evaluator should discuss the fact with the decision maker and try to get at the real source of his/her insecurity. If this fails, the problem should be discussed with the contract decision maker.

If you answered "A" correctly and chose a correct answer from "B" and two from "C" you are doing very well. Congratulations! Please proceed with page 18. If you missed any answer, you should review some or all of pages 11 through 14 before going on.

Problem III: A lack of commitment due to a lack of understanding of the methodology on the part of the decision maker.

A You have been conducting an evaluation of an educational program. The top priority decision maker, Mr. Z, seems to be having a lot of difficulty understanding his role in the evaluation and the evaluative procedures themselves. In the initial stages, Mr. Z asked you "Well, what goals do you think we should evaluate?" He also wanted to know why you couldn't identify the program goals yourself. Throughout the evaluation Mr. Z frequently asked your opinion about aspects of the program, despite your repeated explanations about what the evaluator's role in the decision making process should be. Mr. Z generally seemed to expect to have very little direct involvement in the evaluation until the end when he would read the final evaluation report.

B This decision maker may be the sort of person who dislikes making decisions and would prefer that the evaluator take over this function completely with respect to the evaluation, winding up the process by issuing a verdict on the program. In any case, Mr. Z does not fully understand (or, perhaps, accept) what F-H does and doesn't do.

C This problem should be anticipated. It may be avoided in the following ways.

- 1) The outline of F-H presented to the decision makers at the first meeting with the evaluator should be as straight-

forward as possible. Extraneous matter should be weeded out, and there should be many simple examples included at various stages

- 2) In addition, a clear, simple explanation of the decision maker's and the evaluator's roles in an F-H evaluation should be given. A good source for this is Dr. Larry Benedict's A Practical Guide for Evaluation (1973).
- 3) It should also be mentioned that this is only one form of evaluation and that there are others.

D If the problem arises after the evaluation has been going on for a while, there are a few more steps that might be taken:

- 1) Another decision maker within the program might explain a step in F-H to one who is having problems.
- 2) Go through each step carefully with the decision maker. Avoid leaving materials for him or her to figure out and complete alone.
- 3) If the decision maker persists in asking for the evaluator's opinions, they should be given. The evaluator could explain, however, that these opinions are but one source of data and that the decision maker shouldn't rely upon them too heavily.

E If the suggested solutions to these symptoms do not seem to be working, the evaluator should consider that the decision maker may really have philosophical disagreements with F-H (IV); be impatient with the initial processes (VII); be

insecure (VIII); or have been wrongfully identified from the beginning (X). These possibilities should also be investigated, to the extent that the evaluator has the time to do so.

Review #4

If the decision maker expects you to make decisions for him/her or to somehow judge the worth of the program, it is possible that he/she doesn't fully understand F-H. In the space below, give one way that you, as evaluator, can avoid this problem and one way that you can help solve it.

A. Ways of avoiding this problem include:

1. The outline given to decision makers in the beginning should be as straight-forward as possible and include examples throughout.
2. An explanation of the decision maker's and the evaluator's roles should be given.
3. It should be mentioned that this is only one way of doing evaluations and that there are others.

B. Ways of alleviating this problem include:

1. Soliciting other decision makers in the project to help explain F-H steps to a decision maker having problems.
2. Go through each step of F-H slowly and carefully with the decision maker.
3. Avoid leaving materials for him or her to figure out and complete alone.
4. The evaluator may give his/her opinions if the decision maker persists in asking for them. It should be explained, though, that these opinions are but one source of data and that the decision maker shouldn't rely too heavily upon them.

If you correctly answered one each from "A" and "B," that is very good. Please go on to page 23. If you had any problems, review pages 18 and 19 before going on.

Problem IV: A lack of decision maker commitment due to philosophical disagreements

- A You are an evaluator having difficulties with Mr. Y, a decision maker for a physical fitness program for junior high school students. Mr. Y asks for your opinions about the program frequently. He seems to want you to make all the decisions about the evaluation. He has asked you to interpret data for him which you reported and has even said "I am only doing this because I have to." He also has expressed an interest in other methods of doing evaluations.
- B While Mr. Y may fail to understand F-H, and while he may even be a poor or reluctant decision maker, it is also possible that Mr. Y has philosophical disagreements with F-H that have not been brought out in the open.
- C Again, this is a problem which should be avoided, if possible..
- 1) At the first meeting with individual decision makers, the evaluator should point out the assumptions behind F-H as well as the implications of the purpose "to provide data about an enterprise for decision making."
 - 2) If it seems clear at the outset that one or more decision makers are in opposition to F-H, the evaluator should re-consider carrying out the evaluation. The contract might also be re-negotiated substituting another evaluation model, if everyone including the evaluator is in agreement.

D Once the evaluation is underway, a few other approaches might be tried.

- 1) Try to get the decision maker to go along with F-H by saying something like, "This is only one form of evaluation. If you don't ordinarily carry on systematic data collection, why not give it a try?"
- 2) Consider replacing the decision maker with someone more amenable to F-H. Discuss this with the contract decision maker.
- 3) Consider quitting the job if the decision makers try to get you to drastically alter F-H or to do another kind of evaluation which is unacceptable to you.

E Other solutions that could also be tried are those to problem number III (lack of understanding of F-H); number V (decision maker not being "methodologically oriented"); number VI (status hang-up); number VII (decision maker is impatient); number VIII (decision maker is insecure); and number X (decision maker was wrongfully identified).

Review #5

A. The problem of a decision maker having philosophical disagreements with F-H is more easily avoided than solved. What is one thing you, as evaluator, can do if the decision maker shows strong disagreement in the beginning? Answer in the space below.

B. What is one thing you can do after the evaluation is underway if this disagreement becomes evident and you can't persuade the decision maker to go along with F-H?

- A. Ways of dealing with decision maker disagreement to F-H in the beginning are:
1. Evaluator should point out the assumptions behind F-H at the first meeting with decision makers and discuss them in an attempt to find out whether anyone is opposed to them.
 2. Evaluator should similarly discuss the implications of the purpose "to provide data about an enterprise for decision making." A source for both discussions is the "Evaluation Handbook #1, Defining Evaluation."
 3. If enough decision makers are in opposition to F-H the evaluator should re-consider carrying out the evaluation or should re-negotiate the contract substituting another evaluation model, if that is acceptable to all, including the evaluator.
- B. Ways of dealing with decision maker disagreement once the evaluation is in progress.
1. The evaluator might convince the decision maker to give F-H a try, especially if the decision maker does not normally collect data in a systematic way for decision making.
 2. The evaluator might discuss with the contract decision maker the replacement of this decision maker with one more amenable to F-H.
 3. The evaluator may choose to quit the job if a suitable arrangement can't be agreed upon.

If you answered one from "A" and one from "B" correctly, you are doing very well. Keep up the good work and go on to page 28. If you had any difficulties with any answers, review pages 23 and 24 before continuing.

Problem V: Lack of commitment due to the decision maker not being "methodologically oriented."

A Mr. K, a decision maker for Project TRACK, an educational program, is having difficulties with the evaluation of that program which you are carrying out using F-H. He often tells you, "I'm not good at this sort of thing," and asks you to help him identify his goals and parts for the enterprise, match them, operationalize his goals, and do all the things that the evaluation requires of him. He is not used to thinking about goals--he just likes to operate on a "gut" level. So much of his time on the job is spent dealing with crises, anyway, that he rarely has a chance to become involved in long-range planning. He is sure you can't understand the very real practical constraints that his job places on his time and energies.

B The problem here seems to be that Mr. K is not "methodologically oriented." That is, he is not in the habit, or has gotten out of the habit, of trying to approach problems by means of systematic steps. He needs to try to look at a whole project endeavor in order to best allocate his time, rather than letting himself be tossed randomly from "crisis" to "crisis."

C Before the evaluation, there are a few things an evaluator can do to anticipate the above problem.

- 1) The evaluator can find out what, if anything, is being done in the realm of long-range planning for the project. If none is being carried out, the evaluator should be alert for the above problem.
- 2) The evaluator can ask each decision maker, at the initial meeting, the approximate percentage of time each would estimate he or she spends per week dealing with "crises." For those who indicate a high percentage, the evaluator should be sure to get a firm commitment, no matter how small, of some time each week to be devoted to the evaluation.

D Once the evaluation has begun, the evaluator should do the following.

- 1) Try to stay informed of any crises, major or minor, affecting any of the decision makers. If at all possible, hold them to their time commitments for the evaluation. If this isn't possible, be sure to be clear about how much this time that they are giving up would help them with respect to the evaluation.
- 2) While not emphasizing all the complex steps that make up F-H, try with these decision makers to point out the beauty of the methodology as a whole. Mention that extra time spent on one major step will hurt the evaluation by drawing resources from other steps. Stress the planning behind a worthwhile evaluation and that time spent in planning will certainly save time in the long run, possibly helping to avoid "crises."

3) Ask them to be as patient as possible with steps that they might see as to time-consuming or unproductive.

E Other solutions may include those to Problem III (a lack of understanding of F-H); Problem VII (impatience with initial processes); Problem VIII (decision maker is insecure or under pressure); and/or Problem X (decision maker was wrongfully identified).

Review #VI

- A. If a decision maker is constantly tied up in the day-to-day demands of his or her job, never devoting any time to an examination of the program's long-term operation or goals, it is possible that he or she is not "_____ ."
(Write answer in the space provided.)
- B. Give one way to avoid this problem and one way to help alleviate it when it occurs.

- A. "methodologically oriented."
- B. Ways of avoiding this problem include:
1. The evaluator should find out what, if anything, is being done in the way of long-range planning. If nothing is being done in this area, the evaluator should be alert for the above problem.
 2. For those decision makers who say that they spend a large portion of their time dealing with "crises," the evaluator should make sure he/she obtains a firm minimum commitment of time per week (or month).
- C. Ways of alleviating the problem include:
1. The evaluator should try to hold the decision makers to their time commitments or, at least, to inform them of how the diminished time spent on evaluation activities will affect the data they get back.
 2. Try to emphasize to decision makers how F-H is a functioning whole, and stress the planning that is behind a worthwhile evaluation. Also mention that time spent in planning will save time in the long run, possibly helping to avoid "crises."
 3. Ask the decision makers to be patient with the preliminary steps.

If you answered the questions under "Review #VI" correctly, you are moving along very well. Please go on to "Problem VI." If you had any difficulties with this set of questions, review pages 29 and 30 before going ahead.

Problem VI: Decision maker has a status hang up, i.e., believes he/she is above all this nonsense.

A While applying F-H to the evaluation of a health program, Project GLOW, you discover that you are having problems with one decision maker in particular, Mr. X. You don't seem to be able to gain his trust, and he remains skeptical about your ability to carry out a satisfactory evaluation. Every time you explain a new procedure to him, he doubtfully asks, "What can this be used for?" He seems, in general, unfriendly and rather distant, giving you minimal answers to questions and making very little time available to you for evaluation activities. Mr. X often says something like, "Well, if I have to do all that, what are you going to do?" When he has complaints about the way the evaluation is going, he prefers to go to your superior rather than to face you directly.

B It seems that Mr. X has a status hang-up. While he tries to place himself above you and the evaluation proceedings, it is possible that he is really feeling insecure about the whole thing and has a lot in common with the decision maker in problem VIII. With a female evaluator, it is also possible that the decision maker (male or female) has sexist attitudes which keep him or her from putting forth a strong, sincere effort. A third world evaluator may be experiencing racist or ethnic prejudices.

C These attitudes die hard and may not, in fact, ever totally disappear. It is always easier to blame the evaluator for an incomplete or weak evaluation than to realize that you, as decision maker, did not try your best. A few things the evaluator might do in anticipation of this problem are as follows.

- 1) At the first meeting with each decision maker dress well, appear as confident as possible, and be honest about what you expect from them and what they can expect from you.
- 2) Don't try to "con" the decision makers, especially if they ask any questions about other types of evaluation procedures. Be honest and straight-forward, and have this information at your finger tips.

D As the evaluation progresses, if there are decision makers who seem to have this problem, the evaluator should continue the above approaches and try some of those that follow.

- 1) As much as possible, be an expert in your field. If decision makers ask questions you can't answer, get the answers for them as soon as you can.
- 2) Don't always volunteer to do menial tasks--either related or unrelated to the evaluation, e.g., typing, collating, or Xeroxing reports.
- 3) Be open with them, and try to get them to understand what you are doing during each major step of the evaluation.

E It is also possible that this decision maker doesn't want to know what he's doing wrong (Problem III); is insecure (Problem VIII); or was wrongfully identified (Problem X). Investigate the solutions to these problems if the ones above don't seem to work.

Review #7

A A decision maker who remains aloof from you, skeptical about F-H and resentful of the role he or she must play in it may really be exhibiting a _____ . (Fill in the blank.)

B It isn't usually advisable to try to be too friendly with this type of decision maker, as this may aggravate the problem. Two ways of alleviating it are: (Answer in the space below.)

A status hang-up

b Ways of alleviating the problem include.

1. Be an expert in the field of evaluation as much as possible, but don't try to "con" decision makers if you don't know an answer. Just find someone who does know, and get the information back to the decision makers as soon as possible.
2. Don't always volunteer to do menial tasks, e.g., typing, collating, or Xeroxing reports.
3. Be open with decision makers and try to help them to understand what you are doing during each major step of the evaluation.

If you answered "A" correctly and got two from "B" right, you are doing very well. Keep up the good work, and go on to page 38. If you need to review some of the concepts under "Problem VI," look back at pages 33 and 34 before proceeding.

Problem VII: Decision maker is impatient with initial processes.

A The evaluation of Project LEARN seems to be going smoothly. As evaluator, you are having difficulty with only one decision maker, Mr. J, who is a very busy administrator. Mr. J frequently asks why you can't help out by doing some of the procedures. You always have difficulty making appointments with Mr. J. Also, he tried to submit a goals list that was developed months ago for the funding application, rather than going through the whole goals process from scratch. Before attempting anything Mr. J always asks how long it will take. He seems eager to "get something back" from all this.

B This problem, that of decision maker impatience with the initial processes, is linked with problem #1, a lack of time, and may be connected with others, we well.

C This problem is difficult to avoid because of the nature of F-H and the fact that most decision makers, when they embark upon this type of evaluation, are unfamiliar with what is involved.

1) The best prevention for this problem is the "short-forms" approach mentioned in problem #1. The quickest possible procedure from goal to data collection should be employed so that the decision maker would have information back within, hopefully, two or three weeks. After this, some of the gaps could be filled in to make the abbreviated processes more complete.

- 2) A realistic preliminary overview of the methodology should also help prevent this.

D If the decision maker remains impatient with the process, there isn't too much the evaluator can do.

- 1) Try to stress this over-all approach to evaluation, especially the way, with planning, all the steps fit together to produce a worthwhile product which can, itself, be evaluated.
- 2) Try to stay within the allotted time for each step and don't waste resources.

E The decision maker may be impatient because he has a lack of time (Problem I) or because someone else is pressuring him to produce some evaluation data (Problem VIII). Suggested solutions to these problems may help the evaluator reduce the decision maker's anxiety if the above steps fail.

Review #8

A What is one way that you can avoid a decision maker becoming impatient with the initial processes in F-H? Please answer in the space below.

B What is one way you can cope with this problem, once it has occurred? Answer in the space below.

A Ways to avoid decision maker impatience might include:

1. The use of the "short forms" approach so that a decision maker can get data back as soon as possible and doesn't get bogged down in the initial steps.
2. A realistic overview given decision makers in the beginning should help them see that the first steps, though they do take time, do lead to the reporting of evaluation data.

B Ways to alleviate this problem are:

1. The evaluator can stress the over-all approach to evaluation and the fact that it is systematic and produces an evaluation which can then itself be evaluated.
2. The evaluator should try to stay within the time allotted for each step, especially the initial ones, and not to waste resources.

Please continue with page 42 if you have answered this set of questions successfully. Otherwise, review pages 38 and 39 before going on.

Problem VIII: Decision maker is insecure and/or under pressure.

A Ms. B of Project READ is a busy administrator with little time for you and your evaluation procedures. She tries to discourage you by telling you that "I'm not good at this sort of thing," and "I can't think in those terms," She tries to put you off by being over cooperative and then "forgetting" to do what she'd promised. Basically, Ms. B doesn't trust you and is suspicious of what you are trying to find out about her and Project READ. She isn't at all sure, either, that the other decision makers aren't telling you all about their "made-up" problems with the project. She is nervous about putting any of those "silly" lists--goals, parts, and the rest--in final form, preferring not to commit herself to anything. Once, when she got particularly upset with the whole thing, she snapped, "I'm only doing this because I have to."

B Ms. B is a typical reticent decision maker. All of her behaviors could be the outcome of several different problems, but it seems clear that she is insecure about her role in the evaluation and, at the same time, under a great deal of pressure. This problem is common and may cause one or more problems (I-X) or appear concurrently with them. It is most difficult to pin down, and several strategies might have to be tried before this problem seems to be the likely one.

C The steps mentioned to avoid the previously discussed problems might also help avoid this one. There is one additional suggestion.

- 1) Be as non-threatening as possible. Be particularly careful not to resort to jargon the decision makers won't understand.

D There are a few more things that can be done if this problem arises in the course of the evaluation.

- 1) Choose a "surrogate" and work with him or her, either individually or with the original decision maker.
- 2) Try to lead the decision maker when he or she gets bogged down. Give examples and, perhaps, do some of the steps yourself, but only as a last resort, however.
- 3) Listen to the decision maker. Try to pick up the real cause of the problem. Is it the methodology? The fear of "bad data"? Is it the limited resources? Then, attempt the appropriate "solutions."
- 4) When information has been collected, you, as evaluator, might point out to the decision maker the options he or she has with respect to resulting decisions which might be made. This could also include decisions to collect more data, or different kinds in the future.

E Other problems which decision maker insecurity is likely to foster are Problem II (decision maker not wanting to know what he's/she's doing wrong); Problem VI (status hang-up);

and Problem VII (impatience with initial processes). If this problem is also linked with Problem III (decision maker doesn't understand F-H); Problem IV (philosophical disagreements); Problem V (decision maker not "methodologically oriented"); and/or Problem X (decision maker wrongfully identified), the evaluator should seriously consider asking the contract decision maker to drop this decision maker from the evaluation.

Review #9

A A decision maker who is unsure of his or her role in the evaluation or insecure or under a lot of pressure in his or her job can keep the evaluation from running smoothly. Give one way an evaluator can anticipate this. Please use the space below for your answer.

B What are two ways an evaluator can help solve this problem, once it has occurred?

A Ways this problem may be avoided:

1. The evaluator should try to be as non-threatening as possible.
2. The evaluator should be careful of using jargon that the decision makers won't understand.

B Way this problem may be alleviated:

1. The evaluator can use a "surrogate" who has been approved by the decision maker, to carry out some or all of the steps of F-H.
2. The evaluator can try to help the decision maker if he/she has difficulty with some steps, either by doing some steps for the decision maker or providing more examples where necessary for clarification.
3. Try to discover by talking and listening to the decision maker just what it is about F-H which is causing him/her so much anxiety. Then, attempt a solution, if the problem is among those in this booklet, or try whatever seems appropriate.
4. When the evaluator reports data to the decision maker, he/she can point out some of the decision making options that the data could suggest. The decision maker need not take any of the suggested courses of action, but this may clear up some of the mystery of just what his/her responsibility is.

If your answers are right, very good! Please go on to page 47. Otherwise, you might want to review pages 42 and 43 to clarify the above material before continuing.

Problem IX: Documentation may be dull, dry, or threatening.

A In order to save time while doing an evaluation, you often leave materials with decision makers to work with on their own time. You also sometimes give them materials about the major steps of F-H to study at their leisure. This seems to work fine with all the decision makers except for Mr. S. He oftended complains of having difficulty with the materials and postpones the completion of those activities that you ask him to do. When he does do the requested procedures, he never seems to grasp the problem or to be following directions.

B The above indicates that Mr. S was having difficulties with the documentation of F-H.

C This problem is fairly simple to avoid.

- 1) Make sure before leaving printed materials for decision makers that they understand what is expected of them. Give them a rough estimate of how much time should be involved. Tell them that, if it takes a great deal longer, to wait until you return rather than trying to finish it alone.
- 2) Don't give out printed materials unless they are requested. Do everything verbally, possibly using a tape recorder.
- 3) For each major step, give an introductory seminar which all decision makers should attend. Distribute printed materials and ask them to skim over them and ask questions at this time. This would also provide good feedback for the improvement of the F-H instructional materials.

D If the evaluator has already distributed materials and this problem arises, he or she should carefully go over them with the decision maker and try to isolate the difficulties. If the decision maker still isn't satisfied, perhaps another decision maker from the program who has successfully gone through that particular process could help the first decision maker understand what it is he or she has failed to grasp. This should be done as a last resort, however, since the decision maker with the problem must not feel obliged to adopt another decision maker's goals, parts, etc.

E This problem is most likely to be linked with Problem III (lack of understanding of F-H) and/or Problem V (decision maker isn't "methodologically oriented."). The evaluator should try to apply the solutions to these two problems, if those mentioned above don't bring results.

Review #10

- A The problem of dull documentation is (difficult, somewhat difficult, fairly easy) to avoid. (Underline one.)
- B Give two ways that you might use to avoid this problem.

A fairly easy

B Ways to avoid this problem:

1. Before leaving printed materials with decision makers, the evaluator should be sure that they understand what is expected of them and the amount of time it is likely to take. They should be told that, if they have any difficulties or if the task seems to be taking much too long, that they should get in touch with you or wait until you return, rather than trying to finish it alone.
2. The evaluator might avoid giving out printed materials at all, unless they are requested, choosing to do the procedures verbally, possibly with the use of a tape recorder.
3. An introductory seminar at the beginning of each major step of F-H might be held by the evaluator to familiarize decision makers with the processes. Printed materials could be handed out at the end, and decision makers could skim them and ask questions at this time.

If you have successfully answered this set of questions, please go on to the last problem on page 51. If you had any trouble, review page 47 before proceeding.

Problem X: Decision maker was wrongfully identified by the contract decision maker.

A While evaluating Project JOIN, you have the feeling that there is a problem with one decision maker, Mrs. W. It is hard to identify the source of the problem, but you definitely feel that she is only participating because she has to. Mrs. W does what is asked of her grudgingly and with a minimum of effort. She shows no interest in the evaluation as a whole, and it seems that she couldn't care less whether or not the evaluation provides her with any data that she would not ordinarily be able to obtain.

B This total lack of interest and cooperation might be caused by any of the problems discussed previously. It is also possible, though, that this individual never should have been designated by the contract decision maker to participate in the evaluation, possibly because she has no real commitment to the goals of the project or because her involvement in the project is minimal.

C While it may be necessary to involve decision makers in an evaluation in which they have no real interest, this would not be conducive to the production of the best possible evaluation. The evaluator might avoid this by doing the following.

- 1) At the time the evaluation contract is negotiated, check with the tentatively designated decision makers.

Try to determine how agreeable they are to participating in the evaluation.

- 2) At the initial meeting with decision makers, ask how many expected, before the evaluation contract was negotiated, that they would be substantially involved in the evaluation. For those who had not anticipated this, let them know how much of the evaluation resources have been committed to them. If anyone seems unwilling to match this commitment with a reasonable commitment of time and energy, inform the contract decision maker of this. Ask whether the decision maker can either be persuaded to support the evaluation effort or be replaced.
- 3) For those who had anticipated being involved, try to find out their feelings about F-H, once they have been given the overview. If anyone seems particularly non-committal or indifferent, keep in mind that he or she may be uncooperative in the future.

D When the evaluation is proceeding and the evaluator notices a reluctance to become involved on the part of any decision maker, the evaluator should discuss this with the decision maker to try to find out whether there is really another problem causing it. If so, it should be handled in the appropriate manner. If not, perhaps the contract decision maker should be made aware of the difficulty and asked to consider whether the resources might not be better spent on other existing decision makers or on someone new.

E It is most important to determine whether this problem has occurred in the beginning, because it may well foster Problem I (lack of time); Problem III (lack of understanding of F-H); Problem IV (philosophical disagreements); Problem VI (status hang-up); and/or Problem VII (decision maker impatience). It is much easier to deal with this problem in the beginning, either by having the contract decision maker eliminate this decision maker, by using a "surrogate," or by substituting a new decision maker, than to try to handle all the subsequent problems that this may produce.

Review #11

A If a decision maker is not interested in the evaluation activities at all, it is likely that resources expended on him/her will be wasted. Give two ways that this problem may be avoided.

A Ways to avoid this problem:

1. The evaluator can check with those decision makers who have been designated to help with the evaluation when the contract is being negotiated to try to determine how interested they are in the evaluation.
2. At the initial meeting with decision makers, the evaluator should try to find out how many had expected to be involved in the evaluation. They should be informed of the resources that have been committed to them, and, if they do not choose to contribute some of their own time and energies, this should be reported to the contract decision maker.
3. If anyone seems noncommittal or indifferent to F-H after he/she has read the overview, the evaluator should make a mental note of this and anticipate a possible future lack of cooperation.

If you have answered the above questions correctly, you are finished with this workbook. Good job! There are two summaries which follow. These should be read carefully and any other parts of the workbook reviewed before you take the final review.

If you answered the above question incorrectly please review pages 51 and 52 before going on to the summaries.

Summary of ways to avoid a lack of decision maker cooperation.

1. The evaluator should speak with each designated decision maker, as soon as the contract is negotiated, to explain to them what participation in F-H will entail. A broad outline of F-H may be distributed at this time. This outline should be as straight-forward as possible with extraneous matters weeded out and several simple examples of the various stages in F-H. It should include a clear explanation of the decision maker's and evaluator's roles in an F-H evaluation. A good source for this material is Benedict's A Practical Guide for Evaluation (1973).
2. At the first meeting the evaluator should dress well, appear as confident as possible, and be honest about what is to be expected of the decision makers and what they can expect of the evaluation.
3. Tell the decision makers that F-H is just one form of evaluation and that there are others. If any questions come up about other kinds of evaluation, be honest and straight-forward about answering. Have this information at your finger tips.
4. At the first meeting with decision makers, the evaluator should point out the assumptions behind F-H, as well as the implications of the purpose "to provide data about an enterprise for decision making."
5. The evaluator can try to insure that decision makers feel minimally threatened by an F-H evaluation. The evaluator should stress that even data indicating that certain goals

aren't being achieved are very useful because they can help the program personnel decide whether alternative strategies should be undertaken in order to achieve the goals or whether the goals are unrealistic and should be dropped.

6. Be particularly careful not to use words which may be interpreted by decision makers as "jargon" and which they won't understand.
7. Try to determine how agreeable each decision maker is to participating in the evaluation. Ask whether or not he/she had expected to be substantially involved in the evaluation before the contract was negotiated. Let each know how much of the total evaluation resources have been committed to him/her. If anyone seems unwilling to match this commitment with a reasonable commitment of time and energy, inform the contract decision maker of this. Ask whether the decision maker can either be persuaded to support the evaluation effort or replaced.
8. Try to find out how those decision makers who had anticipated being involved in the evaluation feel about F-H, once they have been given the overview. Make a note of anyone who seems particularly noncommittal or indifferent, as he/she may cause problems in the future.
9. The evaluator should talk with anyone who seems to have a strong negative reaction to F-H. If disagreements can't be resolved, the contract decision maker should be made aware of it.
10. All decision makers who are still participating in the evaluation by this time, should be prepared to give an estimate of

the minimum amount of time which can be guaranteed to be devoted to the evaluation each week. The evaluator should make a note of this commitment for his/her planning of activities.

11. If any decision maker is interested in participating in the evaluation but normally has a very busy schedule, it might be a good idea to identify, from the beginning, someone who shares the decision maker's goals for the program, as much as possible, and can fill in for the decision maker when time is short. This person is known as a "surrogate," and the decision maker's approval must be secured before the surrogate can participate in the evaluation.
12. When holding an initial meeting with each decision maker, the evaluator might ask what he/she feels is the major problem facing the program. Anyone who feels there are no problems might be asked what kinds of data he/she hopes to obtain from the evaluation. If the response indicates that the decision maker is looking for a verification of all the positive qualities that the decision maker assumes the program to have, the evaluator should be aware that this person may not really want an F-H evaluation.
13. At the preliminary meeting, the evaluator should ask what, if anything, is being done in the realm of long-range planning for the project. If none is being carried out, the evaluator should be aware that problems relating to decision makers not being "methodologically oriented" might occur.
14. The evaluator can ask each decision maker, at the initial meeting, the approximate percentage of time each would

estimate he/she spends per week dealing with "crises." For those indicating a high percentage, the evaluator should be sure to get a firm commitment, no matter how small, of some time each week to be devoted to the evaluation.

15. If it seems clear at the outset that one or more decision makers are in opposition to F-H, the evaluator should reconsider carrying out the evaluation. The contract might also be re-negotiated substituting another evaluation model, if everyone, including the evaluator, is in agreement.
16. The "short forms" approach to various steps in F-H may be employed so that decision makers may very quickly get data back and see how F-H can work. The processes can be filled in afterwards so that completeness need not be sacrificed to efficacy.
17. If the evaluator intends to leave printed materials about tasks for decision makers to do on their own, it should be very clear what is expected of them and approximately how much time will be required. They should not feel compelled to continue working more than this time if they find that they don't understand some part of the materials.
18. Printed materials should not be relied upon heavily unless evaluator time is very tight. Some kind of introductory seminar may be held for each major step in F-H to be attended by all decision makers who can skim through them and ask questions at this time, hopefully minimizing confusion.

Summary of general procedures for dealing with a
lack of decision maker cooperation.

1. Observe decision maker behaviors which seem to indicate a lack of cooperation.
2. If the behaviors match those indicated in problems I-X, try to use the suggested solutions to solve the "apparent" problem.
3. If these attempts fail, try to determine what the underlying problem really is, either by talking informally with the decision maker or trying some solutions to related problems listed under Section "E."
4. The evaluator should keep a log documenting the lack of decision maker cooperation, the probable diagnosis of the problem, and his/her attempts to solve it.
5. Some problems may even take a third problem-solving attempt. For instance, the problem of decision maker insecurity (VIII) is often very difficult to pin down, as few decision makers who fear that they aren't doing a good job will admit it to themselves, let alone to a program evaluator.
6. When the evaluator feels that he/she has exhausted all possible solutions or just has no more time to spend on the effort, the contract decision maker should be called in and shown the log documenting the lack of decision maker cooperation and efforts taken to alleviate this. The contract decision maker has the option of talking to the uncooperative decision maker and trying to elicit his/her cooperation; dropping this decision maker from the contract; substituting another decision maker for the uncooperative one;

re-allocating evaluation resources from the uncooperative decision maker to one or more decision makers already identified; or discontinuing the evaluation.

References

- Benedict, L. G. Practical guide for evaluation. Windsor, Conn.: Capitol Region Education Council, 1973.
- Mitchell, V. P. Fortune-Hutchinson and me. University of Massachusetts, mimeograph, 1973.
- Mitchell, V. P. Evaluation handbook #1, defining evaluation. University of Massachusetts, mimeograph, 1975.

Final Review

Write your answers in the space provided.

1. For the three scenarios that follow give: a) the most likely problem causing this lack of cooperation; b) one way of avoiding this problem; c) one way of alleviating it; and d) another possible problem which might be causing these decision maker behaviors.
 - A. You are doing an evaluation of Project JOBS. Mr. J is one of the decision makers of the project and is also an assistant superintendent of the school district. He has worked on the administration of the program for two years and seems to think that it's great. He often points out all the good points of the program to you and seems afraid that you will deliberately search out the bad points. Mr. J seems reluctant to commit himself to any lists--especially the goals list--and always asks what they'll be used for. Mr. J doesn't seem to trust you and, when you mentioned that some decision makers have identified some possible problems with Project JOBS, he discounts them, telling you that these problems just don't exist. Mr. J's main goal is to "document the value" of the program.

B. During the evaluation of Project ADD, a new elementary school mathematics program, you have had repeated difficulties with Mrs. T, one of the teachers. Mrs. T is unfriendly and aloof on a personal level and seems to make the smallest amount of time possible available for the evaluation activities. Occasionally she has tried to go over your head when she had complaints about how the evaluation was going. In general, Mrs. T seems skeptical about your ability to carry out an evaluation at all.

C. While evaluating a junior high school small business program, you have encountered only one decision maker who does not seem to be interested in the progress of the evaluation. In fact, Mr. L definitely seems to be participating in the evaluation only because he has to. He does not seem to want evaluation data for his decision making, and, although he does go through the steps of F-H as you requested, he does so grudgingly and with a minimum of effort.

2. Match each problem from column A with one and only one set of symptoms in column B. There may be some entries in column B with no matching problem.

A

B

- | | |
|---|---|
| 1. decision maker is not "methodologically oriented." | a. wants to edit evaluation reports; doesn't trust you; frequently changes lists; goes over your head. |
| 2. decision maker was wrongfully identified by contract decision maker. | b. doesn't return your calls or answer messages; tries to short-cut steps; has to break appointments. |
| 3. decision maker has a status hang-up. | c. Wants you to identify program goals; has difficulty understanding the decision maker's role; often asks evaluator's opinions; expects to have little direct involvement in the evaluation. |
| 4. documentation is dull, dry, or threatening. | d. "I'm not good at this sort of thing;" little long-range planning going on; "crisis management." |
| 5. decision maker has a philosophical disagreement with F-H. | e. anxious to "get something back" from evaluation. |
| 6. decision makers don't want to know what they're doing wrong. | f. fails to complete printed materials left by evaluator. |
| 7. decision maker has a lack of time. | g. decision maker seems to be stalling; doesn't trust you; believes the program can do no wrong. |
| | h. decision maker gives minimum, grudging cooperation; little interest in the evaluation at all. |
| | i. decision maker asks about other evaluation models; wants you to make a judgment on the program; asks you to interpret data. |

3. What problem is the most common and the most likely to be used as an excuse by decision makers to disguise other reasons for a lack of cooperation?
4. What problem is the most difficult to diagnose and the most likely to underlie several other problems?
5. Give seven things an evaluator can do in the beginning to avoid problems due to a lack of decision maker cooperation:
- (6. Give four general steps, in order, that an evaluator might take to deal with a lack of decision maker cooperation.

Facilitator's Handbook to
Dealing with a Lack of Decision Maker Cooperation

Instructions

Students should be given about an hour to read through "Evaluation Handbook #2, Dealing with a Lack of Decision Maker Cooperation." They should complete the review exercises as they go. They should be allowed to ask questions as they occur during the reading of the workbook, but not during the final review. The final review should not take more than half an hour and should be completed without referring to the handbook.

This booklet contains instructions for correcting the final review, minimum points allowed for competency in this skill, and page references for students who missed certain questions.

1.A

- a) Most likely problem - the decision maker doesn't want to know what he's doing wrong
- b) Ways of avoiding this problem:
 1. Try to discover at the beginning any decision makers who believe there are no problems facing the program and any who want the evaluation to provide only positive data.
 2. The evaluator can try to be minimally threatening to the decision makers and to let them know that negative data can be useful for program improvement.
- c) Way of alleviating this problem:
 1. The evaluator should explain to the decision maker that goals which suggest "documenting the value" or "verifying the worth" of a program assume that the value or worth has already been identified.
 2. The evaluator may discuss the problem with the contract decision maker.
 3. The evaluator might try an immediate break-down of a decision maker goal if most of these have come from the test of completeness. This will tell the evaluator whether the decision maker really can specify the goal or whether he/she has been trying to play games with F-H.
 4. The evaluator should be aware of false excuses about why some data can't or shouldn't be collected. These reasons should be checked with another source.

5. If possible, the contract decision maker should be made aware of any decision maker attempt to edit evaluation reports. In the case of the contract decision maker trying to do this, the evaluator should make it clear that this is not acceptable.
 6. Frequent changes of goals lists or other lists should be examined carefully. If they are mere re-wordings of the originals, the evaluator should discuss the fact with the decision maker and try to get at the real source of his/her insecurity. If this fails, the problem should be discussed with the contract decision maker.
- d) Other possible problems:
1. A lack of understanding of the methodology (III).
 2. The decision maker is not "methodologically oriented" (V).
 3. The decision maker is insecure or under pressure (VIII).
 4. The decision maker had been wrongfully identified from the beginning (X).

1.B

- a) Most likely problem - the decision maker has a status hang-up.
- b) Ways of avoiding the problem:
 1. At the first meeting with each decision maker, the evaluator should dress well, and be honest about what is expected of decision makers and what they can expect from the evaluator and the evaluation.

2. The evaluator shouldn't try to "con" the decision makers, especially if they ask any questions about other types of evaluation procedures. The evaluator should be honest and straight-forward and have this information readily at hand.

c) Ways of alleviating this problem:

1. The evaluator should be an expert in the field of evaluation, as much as possible, without trying to fool decision makers if he/she doesn't know an answer. In that case, the evaluator should find someone who does know the answer and get the information back to decision makers as soon as possible..

2. The evaluator should not always volunteer to do the menial tasks, e.g., typing, collating, or Xeroxing reports.

3. The evaluator should be open with decision makers and try to help them to understand what is being done during each major step of the evaluation.

d) Other possible problems:

1. The decision maker doesn't want to know what she's/ he's doing wrong (III).

2. The decision maker is insecure or under pressure (VIII).

3. The decision maker was wrongfully identified in the beginning (X).

1.C

a) Most likely problem - the decision maker was wrongfully identified by the contract decision maker.

b) Ways of avoiding the problem:

1. The evaluator can check with those decision makers who have been designated to help with the evaluation when the contract is being negotiated to try to determine how interested they are in the evaluation.
2. At the initial meeting with decision makers, the evaluator should try to find out how many had expected to be involved in the evaluation. They should be informed of the resources that have been committed to them, and, if they do not choose to contribute some of their own time and energies, this should be reported to the contract decision maker.
3. If anyone seems noncommittal or indifferent to F-II after he/she has read the overview, the evaluator should make a mental note of this and anticipate a possible future lack of cooperation on the part of this decision maker.

c) Ways to alleviate this problem:

1. The evaluator may discuss reluctance to participate on the part of any decision maker with that person to find out if there might be another problem causing it. If this is the case, that problem should be handled in the most appropriate manner.
2. If that is not the case, perhaps the contract decision maker should be made aware of the difficulty and asked to consider whether the resources might not be better spent on other existing decision makers or on someone new.

d) Other possible problems:

1. The decision maker has a lack of time (I).
 2. The decision maker has a lack of understanding of the methodology (III).
 3. The decision maker has a philosophical disagreement with F-H (IV).
 4. The decision maker has a status hang-up (VI).
 5. The decision maker is impatient with the initial processes of F-H (VII).
-
2. The correct answers are: 1-d; 2-h; 3-a; 4-f; 5-i; 6-g; 7-b.
 3. The most common problem is that the decision maker has a lack of time.
 4. The most difficult problem to diagnose is that of a decision maker who is insecure and/or under pressure.
 5. See the list under "Summary of Ways to Avoid a Lack of Decision Maker Cooperation" at the end of Evaluation Handbook #2.
 6. See the list under "Summary of General Procedures for Dealing with a Lack of Decision Maker Cooperation" at the end of Evaluation Handbook #2.

Competency in this Skill

The questions in the final review are worth the following number of points:

Question #

- (1 (A,B,C) - 4 points each; 1 point for each correct part; 0 for an incorrect answer.
- 2 - 1 point for each correct answer; -1 for each incorrect answer.
- 3 - 2 points for correct answer; 0 for an incorrect answer.
- 4 - 2 points for correct answer; 0 for an incorrect answer.
- 5 - 1 point for each correct answer; 0 for an incorrect answer.
- 6 - 1 point for a correct answer; 0 for an incorrect answer.

Totals

Question #

1	- 12 points
2	- 7 "
3	- 2 "
4	- 2 "
5	- 7 "
6	- 4 "
Total	<hr/> 34 points

The minimum number of points by question for competency in this skill are as follows:

Question #

1	- 8 points
2	- 5 "
3	- 2 "
4	- 2 "
5	- 5 "
6	- 3 "

25 points

Remedial Work

Students who did not earn the minimum number of points on a question should be referred back to pages in the booklet for review. Those page numbers are listed below. Students earning less than 25 points overall should review the entire workbook.

Question #	Page(s)
1.A	11-14
B	33-35
C	51-53
2.1	28
2.2	51
2.3	33
2.4	47
2.5	23
2.6	11
2.7	5
3	5-7

Question #	Page(s)
4	42-44
5	56-59
6	60-61

APPENDIX C

Background Questions

Student Reaction to the Workbook

"Defining Evaluation"

Student Reaction to the Workbook

"Dealing with a Lack of Decision Maker Cooperation"

Background Questions

- 1) I (have, have not) had the Evaluation Methodology I course given at the School of Education. (underline one)

- 2) I am presently (check all that apply)
 - a) a full-time graduate student in (specify department or course of study) _____
 - b) a full-time undergraduate student majoring in (please specify) _____
 - c) working full-time as (specify job) _____

 - d) working part-time as (specify job) _____

 - e) a Continuing Education student
 - f) other (please specify) _____

- 3) I (have, have not) had experience doing educational evaluation. (underline one)

- 4) (Omit if you haven't done educational evaluation before.)
Evaluation approaches I have used are (please specify) _____

- 5) I (am, am not) familiar with the Fortune-Hutchinson Evaluation Methodology. (underline one)

Student Reaction to the Workbook "Defining Evaluation"

1. I found the material in the workbook "Defining Evaluation"

(circle one number in each group)

a) easy to follow difficult to understand

1 2 3 4 5

b) clearly stated confusing or jargony

1 2 3 4 5

c) was comfortably took too long
completed within
the given time

1 2 3 4 5

2. I found the review questions (circle all that apply)

1) too easy

2) too difficult

3) challenging

4) too time-consuming

5) appropriate to the workbook content

6) other (please specify) _____

3. I found the final review (circle all that apply)

1) too easy

2) too difficult

3) challenging

4) too time-consuming

5) appropriate to the workbook content

6) other (please specify) _____

4. Please write any other comments you have about the workbook,
the review questions, and the final review in the space below.

