DOCUMENT RESUME

ED 099 682 95 CE 003 147

TITLE	Handbook for the Evaluation of Career Education Programs (Draft).
INSTITUTION	Development Associates, Inc., Washington, D.C.
SPONS AGENCY	Bureau of Occupational and Adult Education (DHEW/OE), Washington, D.C.
PUB DATE	15 Aug 74
CONTRACT	OEC-0-73-6663
NOTE	150p.
EDRS PRICE	MF-$0.75 HC-$6.60 PLUS POSTAGE
DESCRIPTORS	*Administrator Guides; *Career Education; Data Analysis; Data Collection; Educational Objectives; Educational Research; Evaluation Criteria; *Evaluation Methods; Formative Evaluation; Information Dissemination; Interviews; *Program Evaluation; Question Answer Interviews; Questionnaires; Reports; Research Design; *Research Methodology; Test Reviews; Tests

ABSTRACT

Because of the pragmatic need of school districts to respond to questions posed by those interested in their efforts to implement career education and the equally practical need to further its definition, there exists a demand to assess the results of efforts to implement the concept. The main body of this monograph provides an operational approach to meeting these two needs. It outlines steps which a local school district can follow in conducting an evaluation of its career education effort. A longer term result of following these steps would be the accumulation of information necessary to determine the extent to which the objectives associated with the concept have been achieved; that is, the baseline data for determining whether or not these are individual attributes which should be fostered. If the framework provided is followed and the results are shared with USOE, it is expected that both local needs for assessment and national needs for policy information and further theoretical development will be met. (Author/AJ)

HANDBOOK FOR THE EVALUATION OF CAREER EDUCATION PROGRAMS (DRAFT)

BUREAU OF OCCUPATIONAL AND
ADULT EDUCATION
U.S. OFFICE OF EDUCATION
DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Washington, D.C.
Under Contract No. OEC-0-73-6663

Submitted by:

DEVELOPMENT ASSOCIATES, INC.
1521 New Hampshire Avenue, N.W.
Washington, D.C. 20036

August 15, 1974

Civil Rights Compliance

Discrimination Prohibited - Title VI of the Civil Rights Act of 1964 states:

> "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance."

Therefore, any program or activity receiving financial assistance from the Department of Health, Education, and Welfare must be operated in compliance with this law.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Since 1970 the concept of career education has gained considerable prominence in the eyes of educators and the public at large. Career education legislation has been passed by state legislatures and the U. Congress. Funds have been appropriated at the federal, state, and local levels tc implement career education programs. The U.S. Office of Education has appointed an Associate Commissioner for Career Education.

The concept of career education achieved this prominence because of a pervasive sense among educators and laymen that U.S. education is in need of reform. Career education, which is generically defined as, "the totality of (educational) experiences through which one learns about work," is seen as an important response to that call for reform.

The career education concept has achieved prominence considerably faster than it has achieved precise definition. It is prominent, but it is still in the process of being operationally defined. In some ways it is analogous to the concept of "mental health." Mental health is widely held to be desirable; it has been incorporated into local, state, and federal legislation; public officials have been charged with the responsibility of fostering it; and programs have been funded to help people achieve it. As with career education, however, operationally defining and measuring "mental health" is still the subject of debate and is likely to be for some time to come.

To pursue the analogy one step further, the concept of "mental health" has achieved its current measure of specificity because individuals, at various points, decided to impose a priori but logically defensible definitions. They then sought to evaluate the extent to which treatments designed to achieve these dimensions of mental health were achieved, and to evaluate the ultimate impact on the individual of having achieved them. Similar issues are germane to the concept of career education. It is the judgment of many individuals at all levels who have worked with the concept of career education over the past several years that the time has come to deal with these issues and to systematically refine the concept.

The purpose of this monograph is to report the results of a series of actions that were taken to achieve this objective during the summer of 1974, and to suggest a strategy based on these actions which can carry the process of definition another step forward. It should be made clear at the outset that the decisions reported in the pages which follow are viewed as tentative, but essential, steps. They are tentative in the sense that those who made them do not claim they should not be altered over time. They are essential in the sense that until some decisions were made and tested, the process of operationalizing the concept of career education could not proceed.

An Objectives-Based Definition of Career Education

During the first half of 1974, the Office of Career Education of the U.S. Office of Education engaged in an intensive effort to specify the objectives associated with the concept of career education. The effort used as its base theoretical constructs emanating from the work of industrial psychologists and educational theorists largely beginning in the 1950s. Also considered were several recent attempts to define operationally career education, and the experience of programs which have been in operation over the past several years designed to implement career education in public school settings. More specifically, the following sources provided the basis for the results emanating from those efforts:

- theoretical works pertaining to career development, career maturity, occupational maturity, and individual learning and development;
- federal legislation and policy pertaining to career education;
- state legislation and policy statements regarding career education;
- descriptions of state and federally funded efforts to develop career education curricula and efforts to assess the results of utilizing such curricula;
- descriptions of federal, state, and local efforts to implement career education programs in public school systems in all types of settings in each of the 50 states;
- books, articles, and other publications which dealt with the concept and definition of career education; and
- a series of meetings and conferences with state and local individuals engaged in implementing the concept of career education.

Based on these sources, a set of objectives commonly held to be associated with the career education concept emerged. In all, 33 objectives were identified and grouped into nine major categories or outcome areas. Once developed, the set of objectives was reviewed by individuals of varying backgrounds who were engaged in implementing career education efforts. The objectives were generally found to be acceptable by those individuals.

It should be noted that the set of student objectives does not include statements regarding acquisition of technical skills associated with specific careers or occupational clusters. While providing students with technical skills (e.g., carpentry skills, clerical skills, etc.) should be an aspect of a comprehensive effort in career education, implementation of the career education concept may not currently effect the content of vocational courses or the level of proficiency of the students enrolled. Rather, it is presumed that implementation of career education will currently provide a contextual framework for the courses which may affect the number of students enrolled, may effect a change in the type of student enrolled, and may lead to an increase in the number and type of vocational offerings. The experience of the district in placing exiting students has been included as one measure of the appropriateness and the quality of the specific preparation opportunities being provided young people in the overall career education program.

The results of this process are presented as Table 1. In essence, these objectives are considered to be central to the career education concept and to constitute an objectives-based operational definition. It is recognized that further refinement

of the objectives statements may be necessary. They represent current best thinking in this regard and should be viewed as the starting point towards a realistic operational definition.

It should be noted at this point that much of the literature, and even such federal legislation as Part D of the Vocational Education Act, to the extent that they define career education, combine both objectives associated with the concept and processes presumed to be associated with achieving these objectives. In fact, some definitions of the concept focus entirely on the process dimensions. Specifying a set of educational processes associated with the concept as part of the operational definition was considered and rejected. While this eventually should occur, it was concluded that it would be inappropriate to do so at this time.

What follows on Table 1, then, is but one facet of what should eventually become a full definition of the concept of career education. It is presumed that based on a knowledge of how these objectives are best achieved, the process or treatment dimensions will be added.

TABLE I

CAREER EDUCATION OBJECTIVES IN TERMS OF STUDENT OUTCOME STATEMENTS

I. Students will demonstrate increased self awareness:

A. Students will be able to describe their own current abilities and limitations
B. Students will be able to describe their own current interests and values
C. Students will display positive attitudes toward themselves
D. Students will recognize that social, economic, educational, and cultural forces influence their development

II. Students will demonstrate increased competency in basic academic/vocational skills:*

A. Students will be able to demonstrate generally useful numerical skills
B. Students will be able to demonstrate generally useful communication skills
C. Students will be able to demonstrate generally useful information processing skills
D. Students will be able to demonstrate generally useful decision-making skills
E. Students will be able to demonstrate generally useful interpersonal skills

III. Students will demonstrate increased awareness of work values and possess a desire to engage in paid and/or unpaid work:

A. Students will recognize the bases of various work values
B. Students will possess positive attitudes towards paid and unpaid work

IV. Students will demonstrate increased awareness of and knowledge about work:

A. Students will know the major duties and required abilities of different types of paid and unpaid work

*NOTE: While it is recognized that providing students with career skills (e.g., carpentry, typing, accounting, etc.,) associated with specific vocations is an integral aspect of a comprehensive effort in career education, it was decided to exclude statements regarding achievement of such skills from this list of student outcome because currently a large number of career education efforts do not address this area directly. It should be noted that placement of students has been included

B. Students will know differences in work conditions and life styles associated with different types of paid and unpaid work
C. Students will know entry requirements for major types of paid and unpaid work
D. Students will know the impact of social and technological change on paid and unpaid work
E. Students will know the important factors that affect work success and satisfaction

V. Students will demonstrate increased competency in career decision-making skills:

A. Students will be able to associate their own abilities and limitations with possible success in present or future paid and unpaid work
B. Students will be able to relate their personal interests and values to types of paid and unpaid work and their associated life-styles
C. Students will be able to identify, locate, and utilize sources of information to solve career decision-making problems
D. Students will be able to determine the potential for future advancement/personal growth in work of their choosing
E. Students will know the steps to be taken and the factors to be considered in career planning
F. Students will demonstrate active involvement in career decision-making

VI. Students will demonstrate good work habits:

A. Students will be able to plan work effectively
B. Students will be able to adapt to varied work conditions
C. Students will have a positive attitude towards the concept of quality in relation to a work task
D. Students will have a positive attitude towards conservation of environmental and human resources in accomplishing work tasks
E. Students will have a positive attitude towards responsibility for their own behavior and accomplishment of self-imposed work tasks
F. Students will demonstrate a desire for continuous learning, both in school and out

VII. Students will demonstrate work-seeking and work-getting skills:

A. Students will be able to identify, locate, and utilize sources that contain information about existing paid and unpaid work possibilities
B. Students will be able to demonstrate skills required in applying for and accepting work

VIII. Students who are leaving the formal education system will be successful in being placed in a paid occupation, in further education, or in unpaid work that is consistent with their current career education.

IX. Students will be aware of means available for continued education once they have left the formal educational system:

A. Students will be able to identify sources of additional education in major types of paid and unpaid work
B. Students will be able to identify means to support additional education for themselves in major types of paid and unpaid work

Assessing Career Education

Because of the pragmatic need of school districts to respond to questions posed by those interested in their efforts to implement career education and the equally practical need to further its definition, there currently exists a demand to assess the results of efforts to implement the concept. The main body of this monograph provides an operational approach to meeting these two needs. It outlines steps which a local school district can follow in conducting an evaluation of its career education effort. A longer term result of following these steps would be the accumulation of information necessary to determine the extent to which the objectives associated with the concept have been achieved; that is, the baseline data for determining whether or not these are individual attributes which should be fostered. In addition, the result would provide the data necessary to determine the most effective ways of achieving these objectives.

It is recognized that few if any school systems have explicitly attempted to achieve all of the objectives associated with career education. It is also recognized that resources at the local level are limited. Thus, what is provided is a framework for assessing efforts in career education rather than a tight prescription. If the framework is followed and if the results are shared with USOE, it is expected that both local needs for assessment and national needs for policy information and further theoretical development will be met. As a significant by-product, it is further expected that the administration of career education efforts will also be improved.

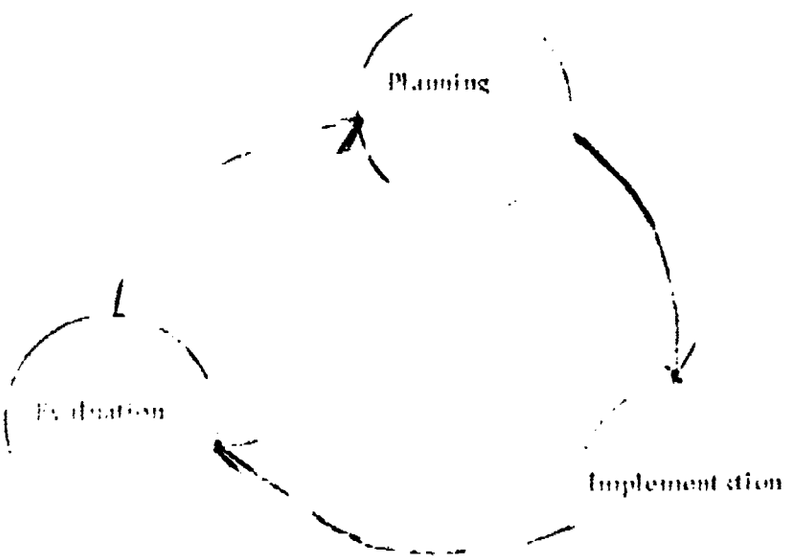
The chapter which follows contains a schematic and narrative overview of a suggested approach to evaluating career education at the local level. The remaining chapters will be elaborations of this overview. In preparing this monograph, it was recognized that levels of experience with educational evaluation vary greatly. As a result, some readers may find it somewhat difficult while others may find it disappointingly simple. An attempt was made to achieve a middle ground with the understanding that some readers may choose to skim some chapers while others may choose to followup on the footnotes and bibliography which provide greater depth.

CHAPTER II

PLAN FOR EVALUATING CAREER EDUCATION

Evaluation is not an end unto itself. Some authors refer to evaluation as a part of the planning process. For purposes of this monograph it will be discussed as part of the management process. This process is viewed as beginning with the development of a plan to reach certain goals or obtain certain results. Next the plan is implemented and then the results are evaluated. Ideally, planning the evaluation should occur after the goals have been set and plans for achieving them laid out, but before those plans have been implemented.

Based upon the results of the evaluation, the original plan is revised or improved as appropriate; the revision is implemented, and its results are evaluated. From this perspective, evaluation is simply the process of examining present performance in order to improve future performance.



to provide assistance in meeting two

al. Two basic assumptions

related guidelines. The primary

would be of assistance to program

rs in formulating evaluation designs.

utilized at the local level and should

primer assumed that each local

of potential value at the national level.

translating specific local objectives

. This translation is necessary in order for

national level to utilize the diverse local

jectives.

ered to be a continuous or cyclic process.

are expected to provide the basis for program

aluation.

evaluation plan presented here is primarily

systems evaluate their efforts in career

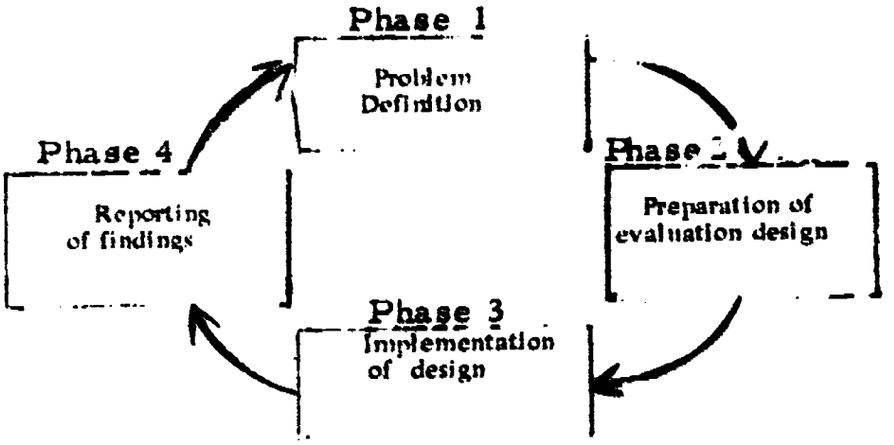
of the plan is to provide information

rs at the local level, it is also designed

with essential information. These two

have led to the separation of

At the broadest level, there are four phases in evaluating any program: specifying what is to be evaluated; designing the evaluation plan; implementing the plan; and reporting the findings.



Within each of these four broad phases of program evaluation there are several steps which should be completed in sequential order. More specifically, the steps which make up the problem definition and design phases define (and therefore must precede) the remainder of the effort. Thus, these initial phases must make clear what is to be evaluated, why it is to be evaluated, how it is to be evaluated, and how the findings are to be reported. Because they are so crucial to the overall process, they are the focus of the major portion of these guidelines.

To provide an overview of the entire process of evaluating career education in the context of a local school system, Figure II-A specifies 13 steps or tasks which should be completed. The first two of these steps may be considered as the pre-design or problem definition phase

(Phase 1, page 13). Steps 3 through 10 are those necessary to complete the evaluation design (Phase 2, page 13) as they specify a procedure to solve the evaluation problem. Steps 11 and 12 relate to implementation of the design (Phase 3, page 13). Step 13 (Phase 4, page 13) involves reporting the findings on which program modifications may be based, leading to renewal of the evaluation process.

OVERVIEW OF STEPS IN EVALUATION PROCESS

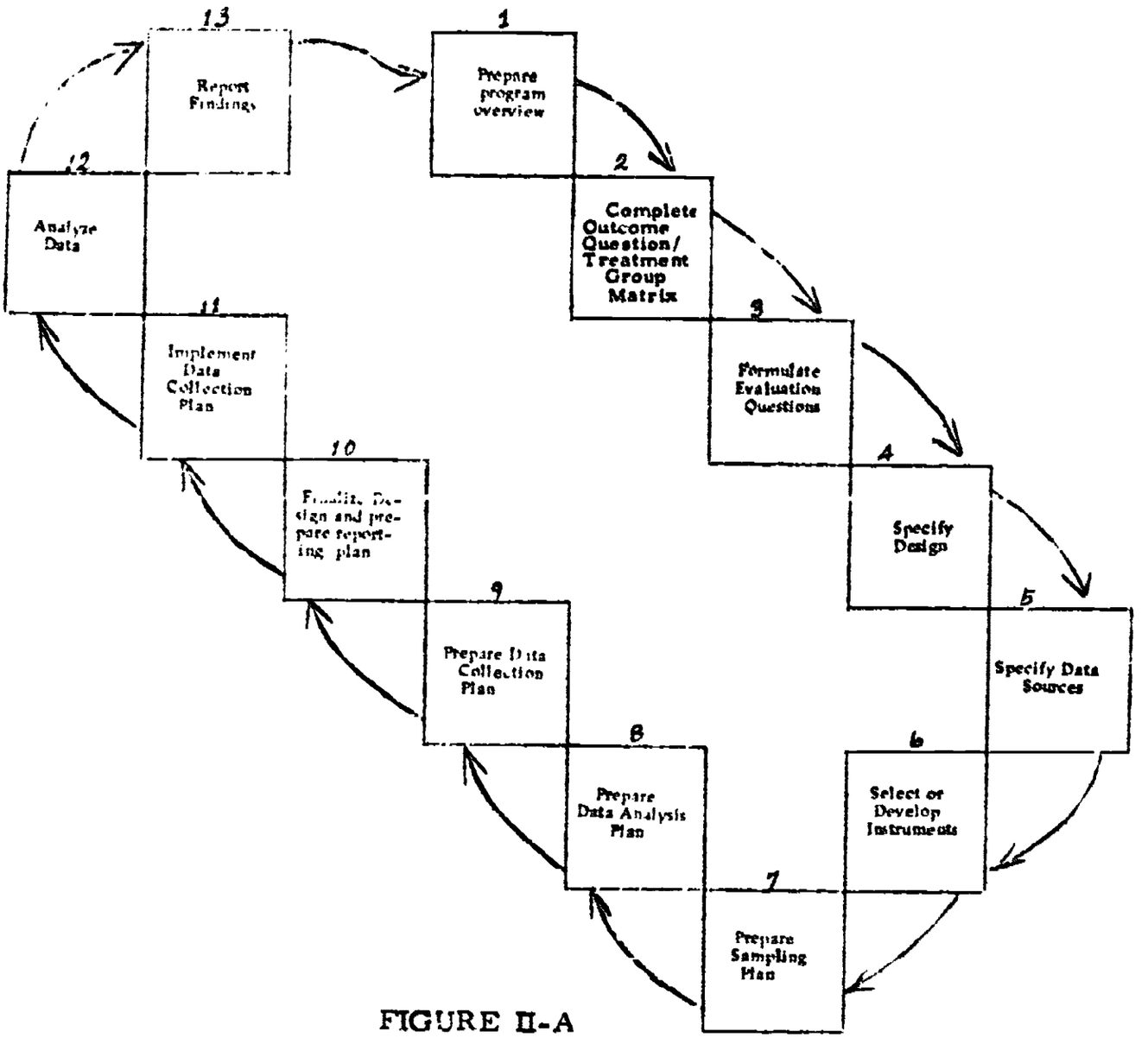


FIGURE II-A

A narrative overview of the 13 steps is provided below:

Phase One - Problem Definition

The first step is to prepare an overview of the program. This defines in general terms what facet of the school system is and is not to be evaluated. It should be completed by school staff as the basis for negotiation with external evaluators or as the basis for determining what external assistance, if any, is needed. Preparation of the overview is the subject of Chapter III.

The second step is to complete a series of tasks leading to completing an Outcome Question/Treatment Group Matrix. In completing these tasks the expected relationship between the student outcome objectives of career education, the activities or treatments which will lead to these outcomes, and the students who receive these treatments in the project are identified. Essentially, this is a further specification of the narrative overview. In some situations. it may be advantageous to involve external evaluators at this point in the process. While the use of external evaluators is not essential,* if such assistance will be used, this is a desirable point at which to include it. Preparation of the matrix is the subject of Chapter IV.

Phase Two - Evaluation Design

The third step is to formulate the evaluation questions. There are three possible types of questions: those pertaining to student outcome objectives, other program objectives, and program management. For example, the

*Some programs, e.g., Part D, require their use.

student . ome type questions might seek descriptions of the extent to which the outcome was attained, descriptions of the extent and quality of treatments, and an assessment of the relationship between treatment and outcome. The basis for formulating student outcome questions was laid in the preceding step; formulating the questions is discussed in Chapter V.

The fourth step is to specify the evaluation design. While this should be implicit in well-formulated questions, it is helpful both to those attempting to relate results of several evaluations, and to clarify thinking at the local level, to make basic decisions explicit regarding the need for pre and post measures and comparison groups. Design considerations are discussed in the context of question formulation in Chapter V.

The fifth step is to specify the data sources to be used to answer the evaluation questions. Some possible sources of data are: those affected by the career education effort(s), those who cause the effects, and observers of the effects. For student outcome questions selection of data sources and data collection instruments are discussed in Chapter VI. For treatment questions data sources and instruments are discussed in Chapter VII.

The sixth step is to select or develop data collection instruments. For several outcome questions and most process or treatment questions, instrumentation must be developed specifically for the evaluation effort. To answer some student outcome questions, a limited number of

instruments are available and are presented for consideration in Chapter VI.

The seventh step is to prepare a sampling plan. The details of the plan depend on decisions made in preceding steps and administrative considerations. The guidelines in Chapter VIII suggest points to consider in making sampling decisions.

The eighth step is to prepare a data analysis plan. While detailed analysis decisions ought not be made until after data is collected, provision for data processing and an outline of the anticipated plan should be prepared at this point. Data analysis is the subject of Chapter IX.

The ninth step is to prepare a data collection plan. Here the tasks, personnel, and target dates for implementing the design are specified. Preparation of a data collection plan is the subject of Chapter X.

The tenth step is to finalize the design and prepare a reporting plan. All decisions made to this point are reviewed in terms of technical feasibility, cost, and the needs of users of the evaluation findings. Once this is done and modifications made, as necessary, a plan for reporting the findings to the various users, along with target dates, is prepared. This process is the subject of Chapter XI.

Phase Three - Implementation

The eleventh step is to implement the data collection plan. It is presumed that problems will arise and all will not go as smoothly as envisioned. It is important to record all deviations from the plan as an aid in analysis and interpretation.

The twelfth step is to analyze the data collected. The first task is processing and tabulating the data. Then suitable analytical techniques are selected and analysis is performed. Steps eleven and twelve, as well as reporting, are the subject of Chapter XII.

Phase Four - Report Findings

The final step is to report the findings and make recommendations as appropriate. The report should be prepared in sections on the assumption that not everyone will have need for all sections of the report. As an aide in comparing the findings of evaluations across projects, the report should provide basic data indicating the context in which the career education effort was carried out.

It should be noted that the emphasis throughout is on obtaining student outcome and related treatment data. While other possible types of evaluation questions are not discussed explicitly, e.g., management processes, the process outlined is equally applicable. Once the question has been formulated, the steps to be accomplished are the same.

B. Design Work Sheets: An Aid to Planning

As an aid in implementing several steps in the evaluation plan, and as a summary of the key decisions which may be of interest to readers of the final report, Table II-A contains an evaluation design work sheet. This work sheet is designed to provide a summary of the evaluation plan in a convenient and logical format. The last step of the evaluation design process described in this manual is to complete this worksheet.

The worksheet is provided here at the beginning of the discussion to provide the reader with a preview of the system. It also provides a handy index and set of instructions for later use, as instructions for completing the work sheet are included in the appropriate columns as well as chapter references. A separate sheet(s) should be completed for each student outcome area.

The reader is asked to briefly review the logic and format of the worksheet at this stage. Each of the following chapters relate specifically to it.

TABLE II-A

EVALUATION DESIGN WORKSHEET

COLUMN 1. Student Outcome Area (See Chapter IV)	COLUMN 2. Related Evaluation Questions (See Chapters IV & V)	COLUMN 3. Design (See Chapter V)	COLUMN 4. Specify Data Source(s) (See Chapter VI & VII)	COLUMN 5. Instrumentation (See Chapters XI & XII)	COLUMN 6. Sampling (See Chapter VIII)	COLUMN 7. Specify Analysis (See Chapters X & XI)
List applicable outcome area. See Outcome Activity Matrix completed in step 2.	Outcome: State outcome question	Indicate decisions regarding pre-post and comparison	List sources from whom data is to be collected	Cite instrument or describe procedure	Specify sampling procedure and size	Specify descriptive and inferential statistics
	Treatment: State related treatment questions	Indicate decisions regarding design	List sources from whom data is to be collected	Describe type of instrument or procedure	Specify sampling procedure and size	
	Relational: State question relating outcome to treatment(s)	Not Applicable	Not Applicable	Not Applicable	Not Applicable	Specify Analysis Techniques
	Outcome:					
	Treatment:					
	Relational:					

CHAPTER III

STEP 1: PREPARING AN OVERVIEW OF THE PROGRAM

To design an evaluation, there must be something to evaluate. Simplistic as this may sound, the first problem facing an evaluator is to determine what there is to study. While with this step it is not necessary to describe the program in great substantive detail, it is essential to establish that something exists which is distinguishable as a discrete subject for the evaluation. The purpose of the overview, therefore, is to convey in simple language to the designer of the evaluation, and ultimately to the reader of the final evaluation report, just what it is that is to be evaluated; that is, to briefly describe the subject of the evaluation and to provide some supporting, background, and contextual material. The overview should be considered as an orientation document for which the evaluation team makes clear what is, and is not, to be included in the evaluation.*

To be most helpful, both in developing the evaluation design and as a draft of a portion of the final report, the overview should address in broad terms the questions of: 1) what are the purposes of the career education effort, 2) what are the major activities undertaken to achieve these purposes, and 3) what resources are used. The questions should be answered for the period to be covered by the evaluation.

*If outside evaluators are to be used, the overview should be included in the materials on which evaluators will base their technical and cost proposals.

If the program is in the planning phase when the evaluation effort is designed, the overview should be based on expected inputs, processes, and program objectives. If the program is on-going, the overview should set forth the expectations for the coming year. Care must be taken to insure that the overview is not based solely on the expectations of prior years. Typically, it is not appropriate to utilize sections of program proposals, annual reports or other previously prepared documents as the overview, although such documents will be a major aid. If there are significant breaks with the past, these should be noted.

It may be helpful to conceive of the overview as an executive summary of the program to be read by the school superintendent, the general public, and interested officials at the national level. In more concrete terms, the overview should contain the following information in summary form:

- Origins of the program. This section will discuss the background of the entire program; how it got started; its original goals; and how it changed over the years.
- Program goals and objectives. This section will state and discuss the present goals and objectives of the program and how they are interpreted by the staff. Program priorities will also be discussed.
- Staff and student participants. This section will discuss the number and types of staff members involved in the program. The responsibilities and duties of each will be summarized.

In addition, the number of student participants will be given by appropriate sub-groups such as grade level, course enrollment, school, etc.

- Major program activities. This section will summarize the major program activities and discuss how the staff and student participants are involved. This will provide an understanding of how the program goals and objectives are being met.
- Amount and source(s) of funds. This will be a listing of incoming funds and major types of expenditures. This will provide an understanding of the size of the program and its emphases.

The overview need not, in fact ought not, be lengthy or require much time in preparation. It should be a terse statement which defines the subject of the evaluation. From the overview the reader should be able to understand whether the entire system is to be included in the evaluation or, if not, what portion is included. The reader should be able to understand whether all efforts which might be considered as contributing to outcomes of students associated with career education are to be considered in the evaluation or whether only selected efforts are to be included and the basis for making that selection. It should be clear to the reader how long the efforts have been going on in the school district and who is responsible for them. Finally, it should be clear who is providing services and who is receiving them.

The importance of clarifying these basic evaluation issues at the start of the process cannot be overstressed. The term career education is subject to multiple definitions. Many of the objectives of career education may overlap with objectives of other programs or emphases of a school district. Activities designed to lead to achievement of career education objectives may be supported financially from several sources and staffed by personnel from various areas in the district.

A recurring problem in the evaluation of career education efforts since 1970 has been to specify what aspect of the school program is to be covered. In some situations it may be appropriate to define the subject of the evaluation in terms of a particular funding source or combination of funding sources. For example, if a school district is supporting career education through funds received from a federal grant, state career education funds, and local revenues, it may be appropriate to include all activities regardless of funding source. What is most important in terms of the evaluation design is that this decision be made explicit at the start of the design process.

Of equal importance to the design is making explicit those broad program activities which will be included in the evaluation. For example, most school systems have taught vocational skills courses for decades. In many districts work experience (cooperative education, distributive education, etc.) courses have been on-going for years. Clearly, vocational skills courses and work

experience programs address many of the objectives of career education programs and thus legitimately might be included in the scope of the evaluation. In some school districts these programs have been added or expanded in recent years as a result of a commitment to career education. In some districts the quantity of effort has remained constant for many years but substantive changes have occurred recently and may be associated with career education. In other districts there have not been changes either in the quantity or the quality of vocational courses or work experience programs but since they are conceptually related to a comprehensive career education program they are defined by the district as part of their career education efforts. In designing the evaluation it is essential that it be specified at the outset whether such activities are to be included in total, included in part, or excluded from consideration.

Closely related to the need to specify which, if any, long-standing efforts of the school district will be included in the scope of the evaluation is the need to specify what new activities are to be included. If the primary focus of the evaluation is in terms of funding sources or use of specific school personnel, then a decision must be made regarding the inclusion of related activities. For example, if a career education coordinator devoted some time to developing a proposal for state funds, should the activities under this grant be included? Similarly, if the evaluation is to focus on a particular program supported by a specific grant, should additional activities also be included in the evaluation that were by-products of the particular program but are not supported by the grant

or are not staffed by individuals directly associated with the grant?

There are no right or wrong answers to the above questions which can be generalized to all situations. What is important is that they be addressed directly at the start of the evaluation process. If they are not, considerable time may be wasted in completion of the succeeding steps and it will be extremely difficult for readers of the evaluation findings to be certain that they understand the basic nature of the effort which was evaluated.

Once the overview of the program is complete, it should be possible to move easily to the next step in the evaluation process and complete a series of tables which make the broad statements of objectives, activities and participants increasingly more specific.

CHAPTER IV

COMPLETION OF OUTCOME QUESTION/TREATMENT GROUP MATRIX

The primary purpose of this framework for evaluation is to assist LEAs in evaluating their efforts in career education. Use of this framework should contribute to the body of career education experience by permitting educators to accumulate experience and compare the effect of different techniques and approaches. The accomplishment of these broad goals requires the translation of local program activities and objectives into sets of common terms which makes such accumulation and comparison possible. While it is recongized that local differences may make such translation difficult, it is important to do so to provide a basis for strengthening the development of career education.

In this chapter a process is described which results in the development of a matrix which describes the basic characteristics of the career education effort being evaluated. In completing the tasks leading to preparation of the matrix, program information will have been organized in a manner which greatly facilitates completing subsequent steps in the evaluation process. Thus, while the matrix itself will be of primary value to those interested in comparing efforts in several locations, and to local people interested in obtaining an overview of the program, completing the tasks leading to its construction are crucial in developing the evaluation design. These tasks may be summarized as follows:

- Specify career education treatments associated with those activity or treatment areas applicable to the local situation.

- Identify the student receiving each treatment.
- Identify groups of participants receiving the same treatment(s). These will be considered the "treatment groups."
- Select the student outcome questions appropriate for each treatment group and complete the Treatment Group-Outcome Area Table.
- Select the more specific student outcome questions appropriate for each treatment group and complete the Outcome Question/Treatment Group Matrix.

Each of these five steps is discussed in somewhat more detail in the pages which follow.

A. Specify Treatments

The product of this task should be a list of the career education treatments or activities provided students. The broad limits of what will be included on this list were established in preparing the program overview (Evaluation Step 1, see Chapter III). For example, if it was determined that the evaluation should encompass all career education activities within the entire school district, then all treatments provided to all students should be included. On the other hand, if the decision was to limit the evaluation to certain types of activities or to those funded by a particular source, then the range of treatments listed here should reflect that basic decision.

In identifying the career education treatments reference should be made to the 12 activity areas listed on the following page in the left hand column of Table IV-A (Activity-Treatment Table). These 12 broad areas were taken from legislative and policy statements regarding career education. They are judged to be on the one hand sufficiently broad to encompass all treatments being provided to students, yet sufficiently distinct to be meaningful as broad categories. The list of career education treatments should be arranged next to these broad headings. Table IV-B provides an illustration of a broad area, elementary familiarization, as it relates to specific, local treatments.

TABLE IV-A
ACTIVITY - TREATMENT TABLE

ACTIVITY AREA	SPECIFIC TREATMENTS
Elementary School Familiarization	
Junior High School Familiarization	
Senior High School Familiarization	
Elementary School Career Orientation	
Junior High Schoc : Career Orientation	
Senior High School Orientation	
Elementary School Career Exploration	
Junior High School Career Exploration	
Senior High School Career Exploration	
Junior High School Work Experience	
Senior High School Work Experience	
Placement (exiting students)	
Other Activities (indicate)	

TABLE IV-B

Illustration

Activity-Treatment Table

Activity Area	Specific Treatment
Elementary Familiarization	1. Curriculum infusion 2. Career education field trips 3. Visitors in the classroom to talk about their jobs 4. Use of special career materials provided by "X" 5. Group sessions with the counselor
Junior High School Familiarization	1. 2. 3. 4. 5.

It is important to consider throughout this task that a career education treatment is by definition something that happens to students which is

distinguishable from other treatments provided by the school system. For example, if field trips are cited as a treatment, they should be a particular type of field trip (defined in terms of purpose, type of advance preparation, or some other factor which distinguishes career education field trips from others).

Several approaches to identifying career education treatments are possible. Given the complexity of most career education efforts, however, experience has shown that a broad-based approach involving planners, administrators, and evaluators is more effective than more limited approaches involving only administrators or only evaluators.

While it may not be possible to describe all of a project's activities in terms of student treatments, every effort should be made to do so. In some cases the relationship between a major project activity and a student treatment may be somewhat indirect but, nonetheless, possible and reasonable to define. For example, the local project may provide in-service training designed to improve an elementary teacher's ability to utilize career education techniques in her regular classroom activities. At first glance one might consider this aspect of the local program to be an "other" activity. Upon careful consideration, however, one would conclude that the in-service training was designed to change the teacher's classroom behavior in order to have a new effect upon students; for example, to familiarize them with the world of work and the range of

opportunities for work. Further, this training for teachers can be expected to result in student outcomes different from the outcomes expected if the training did not occur. Therefore, it is possible to describe this activity as a treatment even though at first glance it did not appear to be.

Although such questions will not be discussed explicitly in this monograph, a whole set of important evaluation questions may be formulated based on the need to assess the effectiveness of such program activities or processes. The process of formulating and answering such questions would be the same as are discussed in this and following chapters.

B. Identify Students by Treatment

Once the list of career education treatments is complete, the next task is to identify the population receiving (or expected to receive) each treatment provided by the program. For each treatment, the participant population should be identified with sufficient specificity to enable an evaluator to identify easily the specific participants for testing or other study purposes.

The ideal method would be to list the names of all participating students by treatment. This is often not practical, however. More feasible may be to group students by grade, course, school, or some other group designation. Experience has shown, for example, that many programs are able to identify all sixth grade students at schools x, y, and z as

receiving a particular treatment. In another program it may be necessary to be more specific. For example, the students in Ms. Jones' third grade class, school X and Mr. Smith's fourth grade class in the same school are those which have a curriculum infused with career education concepts.

The format for reporting the participants by treatment is very straightforward and shown on Table IV-C below. It is simply an extension of the Activity-Treatment Table discussed in Task A.

TABLE IV-C

ACTIVITY-TREATMENTS-PARTICIPANTS TABLE

Activity Area	Treatments *	Participants
Elementary Familiarization	1. Curriculum infusion	Grades 4, 5, and 6 at Brown, Jones, and Smith Elementary Schools
	2. Career education field trips	Grade 4 at Jones Elementary
	3. Visitors in the classroom to talk about their jobs	Grades 1-6 at Brown, Jones, and Smith Elementary Schools
	4. Use of special career materials provided by "X"	Grade 3 at Smith Elementary
	5. Group sessions with the counselor	Grades 5 and 6 at Brown, Jones, and Smith Elementary Schools
Junior High School Familiarization	1. Curriculum infusion	Grade 7 social studies in Main Junior High

*List from Table A.

It should be recognized that in some situations this may be one of the most difficult tasks in the entire evaluation process. Without it, however, it will be impossible to implement the evaluation.

C. Identify Treatment Groups

Based on the table relating treatments and participants, the next task is to identify groups of treatments being received by the same students. The result of this effort will be a list of discrete student groups defined in terms of the manner in which they participate in the career education effort being evaluated. The sum of the number of students in each group should equal the total number of student participants in the program. In other words, care must be taken to insure that the same student does not appear in two groups.

In some situations, experience indicates that very few groups will emerge while in others the list may be quite extensive. To some extent the number of groups will depend on the specificity of the treatments listed in the first task. It is recommended that the level of specificity reflect the type of treatment but not judgments regarding extent or quality. The issues of extent and quality of exposure to various treatments are best handled as questions to be addressed by the evaluation (see Chapter V: Evaluation Questions).

Depending on the situation, completing this task may be a relatively simple or a highly complex undertaking. Essentially, its completion involves inspecting the Activity-Treatment-Participant Table (IV-C) and forming discrete categories of students based on the treatments they receive. Table IV-D which follows illustrates the results of this process based on the information contained in Table IV-C on the previous page.

TABLE V-D
PARTICIPANT - TREATMENT GROUP TABLE

Activity Area	Treatments (See Table IV-C)	Participants (See Table IV-C)	Treatment Groups
Elementary Familiarization	1. Curriculum infusion	Grades 4, 5, and 6 at Brown, Jones, and Smith Elementary Schools.*	A. Grades 5 & 6 at Brown, Jones, & Smith Elementary Schools.*
	2. Career education field trips	Grade 4 at Jones Elementary	B. Grade 4 at Jones Elementary.*
	3. Visitors in the classroom to talk about their jobs.	Grades 1-6 at Brown, Jones, and Smith Elementary Schools	C. Grade 4 at Brown and Smith Elementary.*
	4. Use of special career materials provided by "X"	Grade 3 at Smith Elementary.	D. Grade 3 at Smith Elementary.
	5. Group sessions with the counselor.	Grades 5 and 6 at Brown, Jones, and Smith Elementary Schools.	E. Grade 3 at Brown and Jones Elementary.
			F. Grades 1 & 2, at Brown, Jones, & Smith Elementary.

NOTE: Inspection of the participant column reveals that while all 4th, 5th and 6th graders receive "curriculum infusion," only the Jones 4th graders go on field trips. As a result, the treatment group column reflects the need to differentiate among the 4th grade participants.

The example above was chosen because it represents a slightly more complex configuration than is typically found at the elementary level. In many situations, inspection of the treatment/participant table will yield only two or three groups of students. For example, if the table had indicated that all elementary students went on career education field trips, the number of treatment groups would be reduced from six to four (group B and C would be included in group A, "grades 4, 5, and 6 at Brown, Jones and Smith").

In some situations, particularly at the upper grade levels, it may not be possible at this point in the evaluation process to be certain precisely how students are (or will be) grouped according to specific treatments. In high school, for example, this would require knowing the schedule

of each individual student. In such a case, the most feasible procedure is to estimate, on the basis of informed judgment, how treatments are likely to cluster and then to verify the estimates during data collection and analysis. As a basis for formulating evaluation questions and completing the steps which follow, developing a careful estimate of the number and nature of treatment groups is an extremely important task. In addition, from a program planning and management point of view, the discrepancies between these start-of-the-year estimates and the findings of the evaluation may lead to modification of the effort (or expectations) during the next school year.

D. Complete the Treatment Group -Outcome Area Table

Based on the results of Task C, the product of this task will be a table which relates participants by treatment group, to the broad student outcome area(s) appropriate for that group of students. In Chapter 1 a set of 9 student outcome areas sub-divided into 33 more specific outcome objectives associated with career education were presented. These currently are considered as central to the concept of career education, but are not presented as necessarily exhaustive or in final form. They do, however, represent the current best effort in this regard and as such are the basis for the proposed evaluation approach. The nine broad student outcomes areas presented as statements in Chapter 1 (Table I) are summarized on the following page.

I. Self Awareness

II. Basic Academic/Vocational Skills

III. Awareness of Work Values and Desire to Engage in Paid and/or Unpaid Work

IV. Awareness of and Knowledge about Work

V. Competency in Career Decision-Making Skills

VI. Good Work Habits

VII. Work-Seeking and Work-Getting Skills

VIII. Placement

IX. Awareness of Means for Continued Education

The Treatment Group-Outcome Area Table which is the product of this task consists of a matrix followed by an indication of the outcome area appropriate for each treatment group. An illustration of a completed Treatment Group-Outcome Area Table, based on the information contained in Table IV-D, is presented below:

TABLE IV-E
TREATMENT GROUP-OUTCOME AREA TABLE ILLUSTRATION

Treatment / Treatment Group	Jones Elem. Grades 5/6 at Brown, & Smith Elem.	Smith Elem. Grade 4 at Jones	Grade 4 at Brown & Jones Elem.	Jones Elem. Grade 3 at Smith	Jones & Smith Elem. Grade 3 at Brown	Grades 1&2 at Brown & Smith Elem.
Elementary Familiarization						
1. Curriculum Infusion	X	X	X			
2. Field Trips		X				
3. Visitors	X	X	X	X	X	X
4. Special Materials				X		
5. Group Counseling	X					
Outcome Area	I II	I II	I II	I II	I II	I

From Table IV-E on the preceding page it is possible to see at a glance (a) the various groups of students effected by the career education effort being evaluated, (b) the specific treatments provided each group of students, and (c) the student outcome area(s) expected to be effected by the treatments provided each group of students. Where groups of students receive treatments related to several activity categories, the table should be constructed so as to reflect this situation. Table IV-F provides the basic shell for completing this task for all projects.

TABLE IV-F
TREATMENT GROUP - OUTCOME AREA TABLE
BASIC SHELL

Treatment / Group									
Activity Area:									
1									
2									
• Treatments									
•									
•									
Activity Area:									
1									
2									
• Treatments									
•									
•									
Activity Area:									
1									
2									
• Treatments									
•									
•									
Outcome Area									

E. Complete the Outcome Question/Treatment Group Matrix

On the Outcome Question/Treatment Group Matrix the specific student outcome questions to be answered by the evaluation associated with each broad student outcome area are identified and must be related to the appropriate treatment groups. The completed matrix provides a summary of the specific outcome questions appropriate for discrete groups of students participating in the career education effort being evaluated. This matrix and the Treatment Group-Outcome Area Table on which it is based provide the detailed descriptive information necessary to complete the subsequent steps in the evaluation design. More specifically, taken together they provide the information needed to formulate the final evaluation questions, to develop the sampling plan, and to make initial decisions regarding data analysis.

The matrix consists of rows denoting student outcome questions based on the nine student outcome areas and the 33 related specific objectives presented in Chapter 1. In all, 36 student outcome questions are included as rows of the matrix. The columns of the matrix represent the treatment groups which constitute the local career education effort.

To complete the Outcome Question/Treatment Group Matrix, the column headings must be filled in locally. These will be the same as the headings of the columns of the Treatment Group-Outcome Area Table. Once this is complete, a checkmark denoting the related outcome area should be placed in the appropriate cell of the matrix. Finally, the specific question(s)

for each outcome area should be selected and a mark placed in the appropriate cell.

Utilizing the information provided in Table IV-E (Treatment Group-Outcome Area Table) and decisions regarding objectives, an example of a completed Outcome Question/Treatment Group Matrix is presented below as Table IV-G. In this example, it is presumed that the special materials used by the Smith elementary 3rd graders are specifically related to decision-making and that a change in students' ability to make decisions is an explicit objective only in this facet of the program. The example further indicates that the objective of producing more positive attitudes toward themselves is applicable to all treatment groups and that increasing numerical and communication skills is appropriate for all groups in grades 3 through 6.

TABLE IV-G

Outcome \ Treatment Group	Grades 5 & 6	Jones 4th Grade	Brown & Smith 4th Grade	Smith 3rd Grade	Brown & Jones 3rd Grade	Grades 1 & 2
I.C More positive attitudes toward selves	X	X	X	X	X	X
II.A Increased numerical skills	X	X	X	X	X	
II.B Increased communications skills	X	X	X	X	X	
II.D Increased decision-making skills				X		

Table IV-H which follows presents the shell of the complete Outcome Question/Treatment Group Matrix.

TABLE IV-H

OUTCOME QUESTION/TREATMENT GROUP MATRIX

Outcome Questions \ Treatment Group																
I. Increased Self-Awareness																
A. Have students increased their ability to describe their own current abilities and limitations?																
B. Have students increased their ability to describe their own current interest and values?																
C. Do students display more positive attitudes toward themselves?																
D. Have students increased their recognition that social, economic, educational, and cultural forces influence their development?																
II. Increased Basic Academic Vocational Skills																
A. Have students increased their level of generally useful numerical skills?																
B. Have students increased their level of generally useful communication skills?																
C. Have students increased their level of generally useful information processing skills?																
D. Have students increased their level of generally useful decision-making skills?																
E. Have students increased their level of generally useful interpersonal skills?																

TABLE IV- H-- cont.

Outcome Questions \ Treatment Group													
III. Increased Awareness of Work Values A. Have students increased their recognition of the bases of various work values?													
Increased Desire to Engage in Paid and/or Unpaid Work B. Do students possess more positive attitudes toward paid and unpaid work?													
IV. Increased Awareness of and Knowledge About Work A. Have students increased their knowledge regarding the major duties and required abilities of different types of paid and unpaid work?													
B. Have students increased their knowledge of differences in work conditions and life styles associated with different types of paid and unpaid work?													
C. Have students increased their knowledge of entry requirements for major types of paid and unpaid work?													
D. Have students increased their knowledge of the impact of social and technological change in paid and unpaid work?													
E. Have students increased their knowledge of the important factors that affect work success and satisfaction?													

TABLE IV-H -- cont.

Outcome Question / Treatment Group															
V. Increased Career Decision-Making Skills															
A. Have students increased their ability to associate their own abilities and limitations with possible success in present or future paid and unpaid work?															
B. Have students increased their ability to relate their personal interests and values to types of paid and unpaid work and their associated life-styles?															
C. Have students increased their ability to (a) identify, (b) locate, and (c) utilize sources of information to solve career decision-making problems?															
D. Have students increased their ability to determine the potential for future advancement/personal growth in work of their choosing?															
E. Have students increased their knowledge of the steps to be taken and the factors to be considered in career planning?															
F. Have students increased their active involvement in career decision-making?															

TABLE IV-H --' cont.

Treatment Group / Outcome Questions															
VI. Improved Work Habits															
A. Are students able to plan work effectively?															
B. Are students more adaptable to varied work situations?															
C. Do students have a more positive attitude towards the concepts of quality in relation to a work task?															
D. Do students have a more positive attitude towards conservation of environmental and human resources in accomplishing work tasks?															
E. Do students have a more positive attitude towards responsibility for their own behavior and accomplishment of self-imposed work tasks?															
F. Do students demonstrate an increased desire for continuous learning both in school and out?															
VII. Increased Work Seeking and Work Getting Skills															
A. Have students increased their ability to (a) identify, (b) locate, and (c) utilize sources that contain information about paid and unpaid work?															
B. Have students increased their level of skills required in (a) applying for, and (b) accepting work?															

TABLE IV-H -- cont.

Outcome Questions \ Treatment Group															
VIII. Placement A. How many students have been placed or are engaged in further education and how does this compare with prior years?															
B. How many students have been placed in a paid occupation, and how does this compare with prior years?															
C. Of those placed in (a) further education, and (b) employment, how many consider the placement to be consistent with their career plans?															
D. Of those not placed in further education or in a paid occupation, how many are engaged in (a) unpaid work consistent with their career plans, and how does this compare with prior years?															
IX. Increased Awareness of Means for Continued Education A. Have students increased their ability to identify sources of additional education in major types of paid and unpaid work?															
B. Have students increased their ability to identify means to support additional education for themselves in major types of paid and unpaid work?															

CHAPTER V

FORMULATING THE EVALUATION QUESTIONS AND SPECIFYING THE EVALUATION DESIGN

A comprehensive evaluation of career education in a school district may be expected to address questions dealing with three areas of concern:

- student outcome objectives;
- other program objectives; and
- program management.

For each area, questions may be posed which seek descriptions of some aspect of the program or which seek explanations for descriptive statements. While subsequent discussion will focus on formulating and answering questions related to student outcomes, the same basic considerations apply to the other areas of interest.

In completing Step 2 of the evaluation plan, the questions pertaining to the achievement of student outcomes were identified. In addition, the base was laid for formulating questions regarding the extent and quality of treatment exposure for the students from whom the outcome measures will be obtained. As a result, the base has also been laid for formulating descriptive and explanatory evaluation questions. The formulation of both descriptive and explanatory evaluations questions cannot be separate from a consideration of research design. The remainder of this chapter

therefore discusses the area of design, followed by further discussion of question formulation.

The basic design of the evaluation study is dictated by the specific form of the evaluation questions. The use of relative terms in the question, such as "students exposed to the program will increase their knowledge of..." or "students exposed to the program will become more positive toward..." necessitates the use of a pre/post design, since relational or difference terms of this type require at least two measures, one of which is taken prior to program exposure, or prior to exposure to some significant portion of the program; the other measure would, of course, be taken after such exposure.

Since the career education programs take a significant amount of time to administer, or, to put it in other words, since program exposure must occur over a significant period of time even under the best of circumstances, simply taking pre/post measures of the exposed students and calculating the differences between these measures would not necessarily indicate the effects of the program. This is due to the fact that during the interval between the measures, the students undergo other "treatments" including continuing maturation and the synthesizing of experiences. Thus, a change in the measures cannot logically be attributed to the program.

The usual solution to this design problem is to use a control group; this is to obtain the same pre/post measures, with the same time interval between them, from students who have not been exposed to the program. The simple

logic of this approach of course demands that the non-exposed student be identical to the exposed students in every respect other than program exposure, or that they not be different from them in any systematic way which might affect the results. For example, if the program-exposed students were brighter or less bright than the non-exposed students, or if they were from different socio-economic classes, one could not attribute test score differences between them to the difference in program exposure. The best way to minimize systematic differences between the two groups is to assign students to program exposure and to non-exposure on a random basis. Unfortunately, as is the case with most education programs, this approach is not feasible for career education programs. Not only can students not be assigned randomly, but in many programs it is difficult to assume that non-exposure is possible -- either because the program encompasses all of the students or because teachers and students involved in the program are likely to have an influence on teachers and students who are not actually participating in the program.

Since it is unlikely that a satisfactory control group can be defined in advance, other potentially applicable design approaches should be explored. The most feasible approach is one which selects students, or classroom units, randomly, and obtains information on each student's or each class' exposure to the program. In this way, one can identify the level or "amount" of program exposure for each student or class, and treat those with the lowest

exposure level as a comparison group relative to those with the highest exposure level. Comparisons between the gains made by the high-exposure group and those made by the low exposure group would then provide the basic data needed for assessing whether or not the program had had any impact on the outcome measures, providing that there are meaningful differences between the treatment received by the high-exposure group and those received by the low-exposure group. If the treatment variable is quantifiable on a fairly fine scale, such as number of hours, the "high" group might consist of the 25% of the students who had the greatest number of hours and the "low" group might consist of the 25% of the students who had the fewest hours of exposure.

If the treatment differences are in fact meaningful, the design would lend itself to analysis to determine the strength of the relationship between treatment and outcomes. The framework for this analysis is described below.

The importance of the design in determining the relation between treatment and outcome can be realized by examining the following figure:

Treatments	Outcomes	
	T ≠ C	T = C
T ≠ C	1	2
T = C	3	4

An adequate evaluation of a program is possible for situations falling into cells 1 and 2. In these instances, the treatments will be different for at least two groups, designated as Treatment (T) and Comparison (C), and the outcomes will either be different, as in cell 1, or not different, as in cell 2. (Cell 2, which applies to programs in which the treatments were different but the outcomes were equal would cover situations in which treatment was either inappropriate or not sufficiently

potent to realize the desired outcomes.) Cell 3 would apply to those programs for which the treatments are equal but the outcomes are different. In this case, something apparently was operating to change student or teacher behavior, but specifically what treatments were operating is unknown. Cell 4 includes those programs that show no difference in treatment and no difference in outcomes between Treatment and Comparison groups. This case may be the result of a poorly conceived and/or poorly implemented treatment.

Three questions that should be answered in any treatment-outcome evaluation are as follows:

1. How do the experiences (treatments) for treatment and comparison students differ?
2. How do the outcomes (products) for treatment and comparison students differ?
3. How do the treatments relate to student outcomes (in instances in which there are outcome differences between treatment and comparison students)?

An answer to the first question provides for a decision on which row of Figure 1 is appropriate. If the classroom experiences do not differ for children in treatment and comparison classrooms, then the evaluation need go no further, because any observed outcomes would have no program-related explanation. It is important to emphasize that the differences are not hypothesized differences but actually observed differences in process.

Many a program has planned to implement a certain process, later to discover that the proposed strategies were not implemented, resulting in little, if any, difference between treatment and comparison classrooms.

Given that there are observed differences between treatment and comparison classrooms, the next question is whether outcome differences exist. If outcome differences do not exist, then the process/product aspect of the evaluation need go no further because, although the treatments differed between treatment and comparison, the appropriateness or potency of the treatment was not adequate to influence students' behavior (e.g., self concept, career awareness, etc.). Assuming that there are both treatment and outcome differences between treatment and comparison groups, then the relationship between process and product can be empirically tested (e.g., by using correlational analyses). It is <u>not</u> adequate to simply demonstrate product differences and exposure differences between comparison and treatment groups and then assume that the differences are related. Upon empirically testing the relationship between treatment and product differences, potential causal influences can be identified and later manipulated to test causal hypotheses.

Application of the above outlined approach would provide answers to many of management's questions, including (1) what is the nature of the program, (2) what gains in student outcomes were realized, (3) <u>why</u> did the treatment group gain more than the comparison group, and (4) what process dimensions should be emphasized in subsequent years and what dimensions should be terminated.

From the foregoing it should be clear that the basic evaluation design will be a function of the topical substance of the evaluation question, i.e., does the question require pre and post tests, and the possibilities permitted by the context in which the evaluation takes place, i.e., is a comparison group available. The final formulation of the evaluation questions should reflect the decisions regarding both topical interest and feasibility. For example, while most of the "evaluation" questions posed by USOE imply the need for pre- and post-test measures, they do not address the issue of comparison groups. As a result, final formulation of the outcome questions requires local decisions in this regard.

In practical terms, the situation with regard to comparison groups is likely to resemble one of the following three possibilities:

1. No comparisons possible (all students receive identical treatments; no relational analysis possible).
2. Untreated comparisons (pure control groups, i.e., there is a group of students who are comparable to those receiving career education treatments in all major respects except receiving the treatments).
3. Contaminated comparison (no control group but there is a range of exposure among students in treatments).

Each of these situations is related to a different form of evaluation question. For example, a general question associated with student outcome area I (increased self-awareness) would be formulated as follows under each of the above conditions:

1. Have students exposed to the program increased their ability to describe their own current abilities and limitations?

2. Have students exposed to the program increased their ability to describe their own current abilities and limitations to a greater extent than those students who have not been exposed to the program?
3. Have students exposed to different program treatments increased to an equal extent their ability to describe their own current abilities and limitations?

Primarily as an aid to those attempting to summarize evaluations of career education from several locations, but also as an aid in clarifying the process at the local level, it is advantageous to specify which of the design alternatives was selected. In terms of the design work sheets presented in Chapter II, this simply means recording in column 3 which of the three alternatives with respect to comparison groups is selected. In addition, if for some reason it is impossible to take pre- and post-measures for particular outcome areas, this also should be indicated, and the evaluation question would then be re-formulated to reflect this design change.

CHAPTER VI

MEASURING OUTCOME VARIABLES
SPECIFYING DATA SOURCES AND INSTRUMENTS

Once the evaluation questions have been selected and the design specified, the next step will be to specify for each question the sources from which data are to be collected. The determination of the data sources for each question will be a function of where the data exists, and the availability of existing instruments to gather that data.

This chapter discusses the use of various data sources. The chapter also introduces several data collection instruments that are currently available and attempt to measure career education program outcomes. A table is presented showing the relationship between these instruments and some important student outcomes. This is followed by a brief discussion pertaining to the development of tests locally.

A. Data Sources

It is helpful to classify sources of data relative to student outcomes into two categories: (a) direct and (b) indirect. Direct sources are those which emanate from the students themselves. Indirect sources are those which have their origin in some outside source possessing knowledge concerning student outcome (e.g., teachers, parents, employers, etc.).

In the context of an evaluation effort, direct source data may be divided into two categories. First, that data which must be obtained directly from the students during the course of the evaluation. Second, that data which have already been, or will be, collected from students and stored elsewhere for reasons apart from the evaluation.

1. Direct-source Data

In performing an evaluation which seeks information concerning student outcomes it will be necessary to collect some data from the students themselves. However, in terms of efficiency, it is desirable to maximize data collection from secondary sources. Thus below we first discuss secondary data sources.

a. Secondary Level Data

Even though all outcome data could be sought from the student, it is frequently easier and more efficient to obtain available data from a secondary source. Consider, for example, the situation in which program personnel desire to know how many students have been placed in a work situation during a given period. While that information could eventually be obtained from each student, it would be unnecessary to invest the time and effort in surveying each student if suitable records on such placements were available. It is important, therefore, to first determine what information has been, or will be, generated elsewhere that will be helpful in evaluating the career education effort.

More specifically, the suggested approach to assessing the extent to which students have increased in several dimensions of academic/vocational skills (outcome questions II-A, and II-B) is to utilize the results of the achievement testing program regularly used by the school system. While such tests might not be given on a pre-post test basis, comparisons can be made (a) between participating and non-participating students based on the amount of change between a prior year's test

results and the results of testing during the evaluation period, this on a modified pre-post basis, (b) simply between a group of participants and a matched group of non-participants, a post-only design; or (c) between matched groups of students differentiated by their extent of participation, also a post-only design.

Other information that is normally produced by school systems will also provide indicators useful in the evaluation of a career education program. For example, aptitude test scores, results of interest inventories, and attendance records may all provide meaningful information in assessing the impact of career education upon students. The key to using this kind of information lies in having a clear identification of the program participants. While this may seem obvious, experience has shown that one problem which evaluators frequently experience is that of identifying the treatment population.

In summary, it should not be necessary for the evaluation to duplicate already existing data collection efforts. Existing data should be used before additional data are collected.

b. Primary Level Data

It will not always be possible to obtain direct data from secondary sources; in fact, it will be necessary to obtain most of the information directly from the students. This may be obtained either through the use of instrumentation already developed or through the use of "home made" instruments. As a rule, the utilization of well-constructed and

appropriate instruments is preferable to the development of new and untried instruments. If for no other reason, there results an obvious time saving. A discussion of existing instrumentation and guidelines for the local development of tests of student outcomes is presented later in this chapter.

2. Indirect Data Sources

Not all indicators of student outcomes need be obtained from the students themselves. If, for example, a program decides to use improved school attendance as an indication of improved work habits, the data source might be school or teacher records and not the individual students.

In seeking information concerning student placement, as another example, there are several possible sources of data. One source might be employers. The program might seek to survey area employers to determine how many students have secured employment. A second approach would be to obtain information from students themselves. A mail or telephone survey of students to determine their current employment could be conducted.

Experience indicates that there are problems in obtaining data from these sources. It is difficult to obtain sufficient returns on mail surveys from employers to provide useful information. Followup visits provide better information but require considerable investment of time and money. In all but unusual cases, the costs are prohibitively high.

Attempts to glean placement information from school records (e.g., counselor or placement office) more often than not fail to improve upon the above techniques. School personnel frequently must go from one task or problem to another, leaving the recording of their efforts for later. When time for recording action is available, recall is difficult. The information generated in this manner is usually grossly estimated (not remembered) and of limited value.

Of the three approaches, however, the latter is the one most likely to be used. It is both the least expensive and most managable approach for the school system, since the data source will not be the students themselves, but rather, school records. In designing forms to record information from such records it is necessary to determine what information will satisfactorily provide the answer to the given question, and to consider the nature of the records themselves. In some situations it may be desirable to develop new reporting forms to replace those currently in use. It should be noted, for example, that regular reporting of events and actions generally provides more reliable information than does post-performance data collection. It is one thing, for example, in seeking the number of students placed in full-time employment, to ask personnel "How many students did you place in jobs last year?" and quite another for them to submit weekly data sheets containing this information for the past week.

In securing information concerning student outcome measures which are best obtained from school records or from regular reporting of school personnel, the following procedure should be followed:

- Identify the questions to be answered and the sources of the information (e.g., Question: How many students were placed in part-time paid employment during the school year? Source: school counselor and work experience coordinators).

- Develop a reporting format where one does not already exist.

- Develop reporting procedures as necessary (e.g., weekly report to Career Education coordinator).

- Implement procedures at beginning of period to be evaluated.

Following these procedures will not only insure more reliable information for purposes of evaluation, but also has the potential of producing some useful management information for program directors. The latter benefits would be enhanced if the reporting format would also provide for noting various demographic facts about each student, such as age, sex, and ethnicity. This would also be of value in analyzing other outcome findings and in studying the relation between treatments and outcomes, since the program may be having different effects on different demographic groups.

B. Instrumentation

1. Available Instruments

Referring back to direct data sources, there are several data collection instruments which attempt to measure student career education outcomes. During the summer of 1974, USOE undertook a review of these instruments. While the review was not exhaustive, it was extensive and is presently continuing. A review panel was convened for the purpose of reviewing instruments that might be appropriate measures of career education student outcomes.

In preparation for the panel's work relevant instruments were sought from as many sources as possible. The search for instruments included a review of the evaluations of all first round Part D programs and other career education programs. A survey of test publishers was also conducted and selected third party evaluators were asked to suggest instruments that had proved to be of value in their past efforts. In addition to these sources, local and federal personnel engaged in the operation of career programs were asked to make recommendations of instruments to be reviewed by the panel.

The search produced 90 instruments for consideration by the panel. The tables which follow display the results of the panel's efforts and should be of value in designing an evaluation of a career education program.

The first table (Table VI-A) is concerned with instruments reviewed by the panel and recommended as appropriate for measuring the student outcomes of career education efforts. Table VI-B lists the career education areas for which no tests were recommended.

The third table in this section (Table VI-C) provides specific information for each of the recommended instruments. The last table lists tests which were found to have some potential in measuring student outcomes but were not recommended because of insufficient psychometric data.

Table VI-A which follows, entitled "Instruments Recommended by Panel," shows the relation between the various parts of the recommended instruments and various student outcomes. The table is sub-divided into three grade level groups, K-3, 4-6, and 8-12; applicability for Grade 7 is noted separately where appropriate.

As the remarks section of Table VI-A (and the summary matrix shown in Figure VI-A indicates, most of the subtests or parts of the recommended instruments were found by the panel to be measuring more than one outcome. This means that the scores obtained in such a subtest cannot be interpreted unambiguously. If, for example, there is a gain in the scores for a subtest which measures outcomes A and B, it is impossible to tell whether this represents a gain in outcome A, a gain in outcome B, or a gain in both.

In instances where the multiple outcomes measured by a subtest can be viewed as meaningful parts of some higher level composite concept, the subtest may be viewed as measuring the composite. By way of analogy, an instrument judged to measure length, width, and height simultaneously would be judged to be measuring volume. Since volume is a composite of these three factors, and a change in any one, any two, or all three would necessarily result in a change in volume.

In theory, a gain in the scores for a subtest which measures two (or more) outcomes may also result from a gain in one outcome and a loss in the other (or others). The obvious implication of these points is that preference should be given to those subtests which measure only one outcome, followed by those which measure two outcomes that are clearly related to each other.

TABLE VI-A
INSTRUMENTS RECOMMENDED BY PANEL*

Outcome	Instrument	Remarks
Grades K - 3		
I C. Positive attitudes toward selves	SOS, Primary Level, scales 1, 2, and 4 combined	See Instrument List
IV A. Duties and abilities required	CEQ, Grade K-3	Test assesses IVA and IVB in combination
IV B. Work conditions	CEQ, Grade K-3	
VI F. Desire for continuous learning	SOS, Primary Level, scale 3	See Instrument List
Grades 4 - 6		
I C. Positive attitudes toward selves	SOS, Intermediate Level, scales 1, 2, 3, 4, and 7 combined	See Instrument List
IV A. Duties and abilities required	CEQ, Grade 4-6	Test assesses IV A, IV B, and IV C in combination
IV B. Work conditions	CEQ, Grade 4-6	
IV C. Entry requirements	CEQ, Grade 4-6	
VI F. Desire for continuous learning	SOS, Intermediate Level, scale 5	See Instrument List
Grades 8 - 12		
I C. Positive attitudes toward selves	ACD, Unit 4, Items 37 and 38	----------
I D. Recognition of forces influencing development	CDI, Scale C	Measures I D, IV A, IV C, V C, and V E in combination
III B. Positive attitudes toward work	CMI, Attitude Scale	Also applicable for Grade 7
IV A. Duties and abilities required	• ACD, subscore 1	Measures IV A and IV B in combination
	• ACD, subscore 3	Measures IV A and V B in combination
	• CDI, scale C	Measures I D, IV A, IV C, V C, and V E in combination
	• CMI, Part I	• Measures IV A, IV E, V A, V B, and V E in combination • Also applicable for Grade 7
	• CMI, Part 2	Also applicable for Grade 7
	• CEQ, Grade 7-9	• Measures IV A, IV B, IV C, IV D, V E, and VII A in combination • Applicable for Grades 7, 8, and 9 only

*Descriptive information about each test is contained in Table VI-C (Instrument List).

Grades 8 - 12 -- cont.

Outcome	Instrument	Remarks
IV C. Work requirements	• ACD, subscore 1	Measures IV A and IV B in combination
	• CEQ, Grade 7-9	• Measures IV A, IV B, IV C, IV D, V A, and VII A in combination • Applicable for Grades 7, 8, and 9 only
IV C. Entry requirements	• ACD, subscore 2	
	• CDI, scale C	Measures I D, IV A, IV C, V C, and V E in combination
	• CMI, Part 4	Also applicable for Grade 7
	• CEQ, Grade 7-9	• Measures IV A, IV B, IV C, IV D, V A, and VII A in combination • Applicable for Grades 7, 8, and 9 only
IV D. Impact of social and technological change	• ACD, subscore 4	Measures IV D, V C, and V E in combination
	• CEQ, Grade 7 - 9	• Measures IV A, IV B, IV C, IV D, V A, and VII A in combination • Applicable for Grades 7, 8, and 9 only
IV E. Factors affecting success and satisfaction	CMI, Part 1	• Measures IV A, IV E, V A, V B, and V E in con bination • Also applicable for Grade 7
V A. Relating own abilities to work possibilities	• CMI, Part 1*	• Measures IV A, IV E, V A, V B, and V E in combination • Also applicable for Grade 7
	• CMI, Part 3*	• Measures V A and V B in combination • Also applicable for Grade 7
	• DAT	• See Instrument List regarding scoring.
V B. Relating own interests and values to work possibilities	• ACD, subscore 3	Measures IV A and V B in combination
	• CMI, Part 1*	• Measures IV A, IV E, V A, V B, and V E in combination • Also applicable for Grade 7
	• CMI, Part 3*	• Measures V A and V B in combination • Also applicable for Grade 7
	• DAT	• See Instrument List regarding scoring.

*These sub-tests are based upon the assumption that the ability to relate the skills and interests of others to work possibilities is highly correlated with the ability to do this for oneself.

Grades 8 - 12 -- cont.

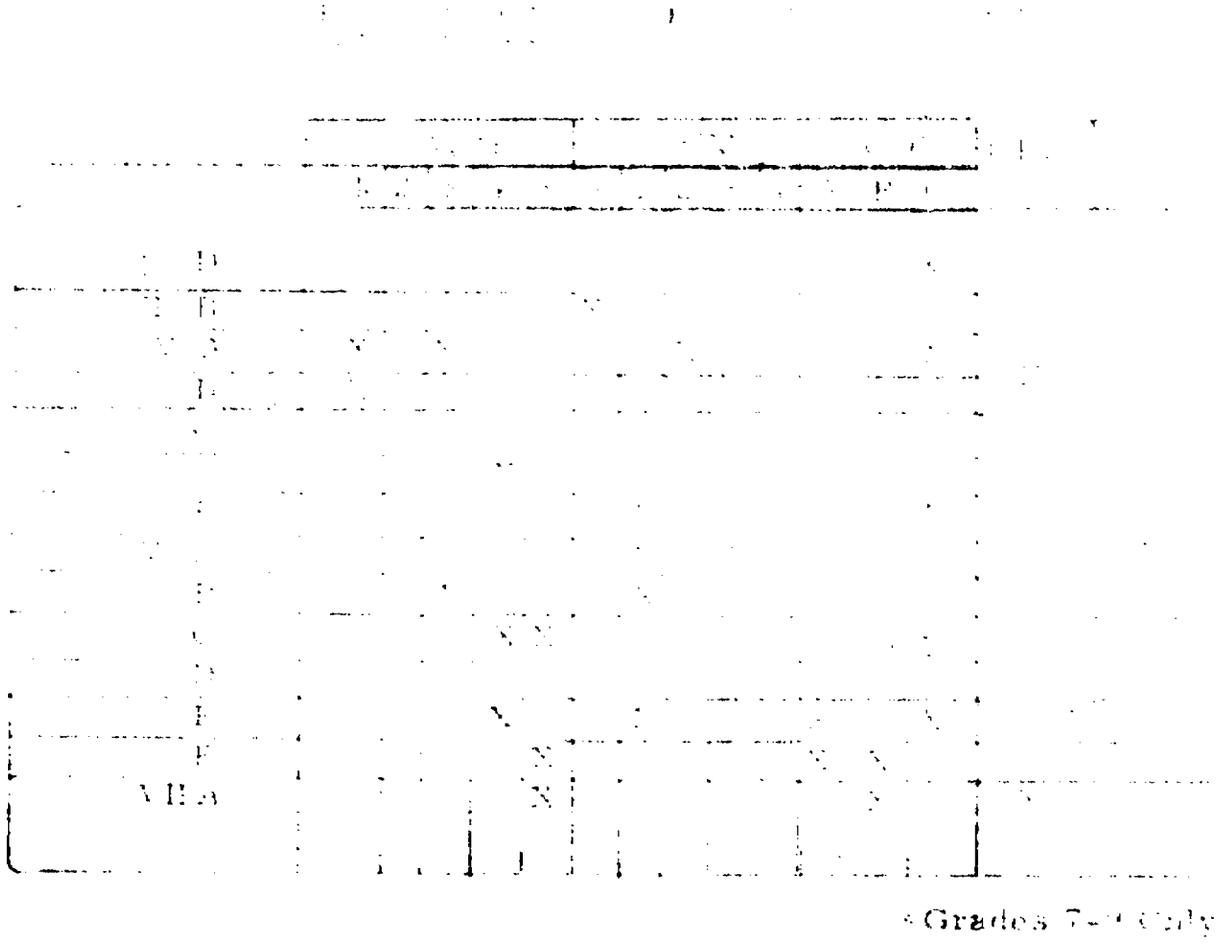
Outcome	Instrument	Remarks
V C. Awareness and use of information sources	• ACD, subscore 4	Measures IV D, V C, and V E in combination
	• ACD, subscore 5	Measures V C, V F, and VII A in combination
	• CDI, scale B	Measures V C, V F, and VII A in combination
	• CDI, scale C	Measures I D, IV A, IV C, V C, and V E in combination
V E. Steps and factors in career planning	• ACD, subscore 4	Measures IV D, V C, and V E in combination
	• CDI, scale A	Measures V E and V F in combination
	• CDI, scale C	Measures I D, IV A, IV C, V C, and V E in combination
	• CMI, Part I	Measures IV A, IV E, V A, V B, and V E in combination
	• CEQ, Grade 7-9	• Measures IV A, IV B, IV C, IV D, V E, and VII A in combination • Applicable for Grades 7, 8, and 9 only
V F. Active involvement in career decision-making	• ACD, subscore 5	Measures V C, V F, and VII A in combination
	• CDI, scale A	Measures V E and V F in combination
	• CDI, scale B	Measures V C, V F, and VII A in combination
VII A. Awareness and use of information sources	• ACD, subscore 5	Measures V C, V F, and VII A in combination
	• CDI, scale B	Measures V C, V F, and VII A in combination
	• CEQ, Grade 7-9	• Measures IV A, IV B, IV C, IV D, V E, and VII A in combination • Applicable for Grades 7, 8, and 9 only

TABLE VI-B

Outcome Areas For Which No Tests Were Recommended
(Grades 8 - 12)

I.	A.	Describing own abilities and limitations
	B.	Describing own interests and values
II.	A.	Numerical skills *
	B.	Communication skills *
	C.	Information processing skills
	D.	Decision-Making skills
	E.	Interpersonal skills
III.	A.	Recognition of bases of work values
V.	D.	Determining potential in preferred work areas
VI.	A.	Effective work planning
	B.	Adaptability in work situations
	C.	Positive attitudes toward work quality
	D.	Positive attitudes toward resource conservation
	E.	Positive attitudes toward responsibility
	F.	Desire for continuous learning
VII.	B.	Skills for applying for and accepting work
VIII.	A.	Placement in further education
	B.	Placement in work
	C.	Placement/plans consistency
	D.	Unpaid work/plans consistency (unplaced students)
IX.	A.	Identifying sources of career-related education
	B.	Identifying means to support career-related education

* Since it was assumed assessment of these outcomes would utilize the results of the school districts' existing testing program, instruments pertaining to the outcomes were not included among those reviewed.



* Grades 7-12 Only

FIGURE VI - A

TABLE VI-C -- INSTRUMENT LIST

THE ASSESSMENT OF CAREER DEVELOPMENT (ACD)

Published by Houghton Mifflin Sales Office: Measurement Guidance Dept.
Houghton Mifflin
One Beacon St.
Boston, Massachusetts 02107

Grade levels: 8 - 11

Administration time: 125 minutes (three sessions)

Cost: Test booklets (package of 35) = $ 13.95
Answer sheets (package of 35) = $ 3.75

Scoring: Cannot be scored locally except by special arrangement; summary data for five subscales and 42 additional items are reported.

Remarks:

a. Answer sheet provides for collecting responses for up to 19 locally constructed questions.

b. Should not be used below 8th grade except for groups having above average verbal ability.

c. Subscore key

Subscore 1 - Occupational characteristics
Subscore 2 - Occupational preparation requirements
Subscore 3 - Exploratory occupational experiences, general
Separate scores for each of the following clusters:
- Social/health, and Personal Services
- Business Sales and Management
- Business Operations
- Technologies and Trades
- Natural, Social, and Medical Sciences
- Creative and Applied Arts

Subscore 4 - Career Planning Knowledge
Subscore 5 - Career Planning Involvement

CAREER DEVELOPMENT INVENTORY (CDI)

By Donald E. Super, et al, Teachers College, Columbia University, New York, N.Y., 10027

Available from authors; to be used with their permission.

Grade levels: 8 - 12

Administration time : 30 minutes for average student; all students should complete inventory in one class period.

Scoring: Can be scored locally or commercially.

Remarks:

a. There is sex stereotyping in a few items, but empirical studies show there is no sex bias; revision to remove the stereotyping is in process.

b. Scale key: Scale A - Planning Orientation
Scale B - Resources for Exploration
Scale C - Information and Decision-Making

CAREER EDUCATION QUESTIONNAIRE (CEQ)

Published by: Minnesota Research Coordinating Unit for Vocational Education,
University of Minnesota, Minneapolis, Minnesota

Grade levels: K-3, 4-6, 7-9

Administration time: 30 - 45 minutes

Scoring: Can be scored locally.

Remarks:

a. Each grade level test yields one overall score.

b. Quality of pictures in the K-3 form is not uniformly high; vocabulary level may be high.

TABLE VI-C -- cont.

CAREER MATURITY INVENTORY (CMI)

Published by: CTB McGraw Hill, Del Monte Research Park, Monterey, California 93940

Grade levels: 7 - 12; items should be presented orally to those students who are not able to read at the sixth grade level

Administration time: 2 1 2 hours

Cost: Test booklets (Package of 35) $20.00
Answer sheets (Package of 35) $ 4.00

Scoring: Can be scored locally or commercially

Remarks:

a. The Attitude Scale items may permit the development of meaningful subscores.

b. Minor degree of sex stereotyping; no clear evidence of sex bias.

c. Part Key

A - Attitude Scale

Part 1 - Knowing Yourself (Self-Appraisal)
Part 2 - Knowing About Jobs (Occupational Information)
Part 3 - Choosing a Job (Goal Selection)
Part 4 - Looking Ahead (Planning)
Part 5 - What Should They Do? (Problem Solving)

(Part 5 was not recommended by the review panel due to disagreement regarding the scoring of some of the items)

DIFFERENTIAL APTITUDE TESTS with Career Planning Program (DAT)

Published by: The Psychological Corporation, 304 East 45th Street, New York, N.Y. 10017

Grade levels: 8 - 12

Administration time: 235 minutes

Cost: Test booklets (package of 25) $18.50
Answer sheets (package of 50) $12.50 (IBM and Op Scan)
Answer sheets (package of 50) $11.00 (NCR)

Scoring: Can be scored locally or commercially.

Remarks:

a. Career Planning Program is currently formulated on an individual student basis; publisher is prepared to offer a group summary of findings; the model for this summary is shown in Appendix D.

b. Might be utilized in lieu of achievement testing for outcome II A and II B.

TABLE VI-C -- cont.

SELF OBSERVATION SCALES (SOS)

Published by: National Testing Service, 2626 Erwin Road, Durham, North Carolina 27705

Grade levels: K-3, 4-6

Administration time: 20-25 minutes

Cost: Test booklets and Answer sheets (package of 30) = $35.00

Scoring: Cannot be scored locally.

Remarks:

a. This test was recommended by the review panel only on the condition that the publisher provide users with information regarding the scoring keys and weights for the subscales. (See Appendix C for publisher's comments.

b. Local review for sensitive items was recommended by the review panel.

c. The review panel recommended further work by the publisher on validation.

d. Subscale Key:

K-3 (Primary Level)

Subscale 1 - Self Acceptance
Subscale 2 - Social Maturity
Subscale 3 - School Affiliation
Subscale 4 - Self Security
Subscale 5 - Achievement Motivation
(Subscale 5 was not recommended by the review panel due to insufficient psychometric data.)

4-6 (Intermediate Level)

Subscale 1 - Self Acceptance
Subscale 2 - Self Security
Subscale 3 - Social Maturity
Subscale 4 - Social Confidence
Subscale 5 - School Affiliation
Subscale 6 - Teacher Affiliation
Subscale 7 - Peer Affiliation
Subscale 8 - Achievement Motivation
(Subscale 8 was not recommended by the review panel due to insufficient psychometric data.)

TABLE VI-D

TESTS HAVING PROMISING APPROACHES BUT INSUFFICIENT DEVELOPMENT OR LACK OF PSYCHOMETRIC DATA

New Mexico Career Education Test Series:

Monitor, P. O. Box 2337
Hollywood, California 90028

Orientation of Career Concepts:

Evaluative Research Associates, Inc.
8444 Floussant
St. Louis, Missouri 63121

Affective Assessment Questionnaire:

Minnesota Research Coordinating Unit for Vocational Education
University of Minnesota
Minneapolis, Minnesota

Decision Making Scale:

IBEX
2526 Erwin Road
Durham, North Carolina 27705

PSE Career Education Test:

Policy Studies in Education
52 Vanderbilt Avenue
New York, New York 10017

Attitude Toward Self and Others Scale:

Center for Occupational Education
North Carolina State University at Raleigh
Raleigh, North Carolina

Career Awareness Inventory:

Cornell Institute for Research and Development in Occupational Education
State University of New York
Ithaca, New York

TABLE VI-D (Continued)

J. V. M. -Scale:

School District of the City of Royal Oak
4000 Crooks Road
Royal Oak, Michigan

2. Locally Developed Instruments

For outcomes not measured by these instruments and in planning to secure information from other indirect sources such as teachers, counselors, parents, employers, etc., it will be necessary to consider the development of instruments locally. If it is decided that the need for one or more "home made" tests justifies the time and costs involved in developing such tests, there are several important technical factors which must be considered. Some of these factors apply to each item to be constructed; others apply to the "test" as a whole. In either case, the major considerations pertain to the need to maximize the validity and the reliability of the students' responses, and ultimately of the score or scores which will be generated by the test.

The starting point for maximizing validity is a clear delineation of the subject matter, or outcome area, which the test is supposed to measure, followed by a listing of specific content areas which the participating teachers plan to cover, or have covered, within the outcome area. The test items would then be constructed around the specific facts or concepts which represent the more important teaching objectives.

Having specified the contents of the test items, a decision is needed concerning the format to be used for the items themselves. There is usually a choice among several alternatives, such as True/False, multiple choice, Agree/Disagree, and matching, among others. It is more difficult to construct items in some formats than in others, but the

effort may be worth while. For example, multiple choice items are more difficult to construct then true/false items, but the results are apt to be more stable and more valid since there is a 50/50 chance of getting a true/false item right simply by chance. (Where there is a high probability of being correct on the basis of choice alone, test makers generally construct a relatively large number of items as a way of compensating for chance effects and insuring an adequate spread in the scores achieved by the respondents.) In addition to the ease or difficulty of constructing the items, the choice of format should also be based upon such factors as: the suitability of the format for the intended respondents, the length of time to administer the test using each format, and ease of scoring.

In the course of constructing the items for the selected format, great care must be taken in the wording of each item, and in the case of multiple choice items, in the wording of the response alternatives, to insure that there is no ambiguity and to insure that the vocabulary and sentence structure are within the capabilities of the potential respondents. The language of the item, as well as the test formats, must be so designed as to insure that incorrect responses occur only because the student does not know the correct answer, rather than as a result of difficulty in understanding the item or confusion with regard to the test directions. Both of these factors may be checked by means of a small-scale pre-test of the instrument involving 20 to 30 students.

[illegible] that the group includes as wide a range of ages and abilities as [illegible] respondents.

In the course of revising items to insure that the language is clear and [illegible], it is easy to lose sight of the specific knowledge which the [illegible] was intended to assess; often one finds that the item is easy to understand but does not have direct relevance to the outcome area. [illegible] for the common occurrence may be minimized by having cognizant [illegible] review each item for appropriateness of both the substance, [illegible] of the items, and the specific language and format being used. [illegible] should also review the test as a whole for such defects as repetition of content areas, inadvertently providing the answer to an item in another item, and under-sampling or over-sampling of the priority facts and concepts involved in the teachers' treatments.

In summary, the development of adequate tests is a complex, time consuming undertaking. While there clearly is a need for instruments which adequately address the student outcomes associated with career [illegible], development of tests should be attempted only with great [illegible].

In the chapter which follows, several other types of instruments which [illegible] be developed locally are discussed. While the discussion will be in the context of measuring treatment variables, some of the techniques [illegible] the discussion of questionnaires) may be useful and easily adapted to the context of measuring student outcomes.

CHAPTER VII

MEASURING PROGRAM TREATMENTS

Process or treatment evaluation is important because program managers need to know "why" a program worked just as much as they need to know how successful it was. In general terms, accumulation of data required to measure project treatments is based on the same series of steps taken in measuring the student outcomes. Specifically, it is first necessary to identify treatments and to formulate a set of treatment questions. Second, it is necessary to specify those who will provide the information necessary to answer the treatment questions. Third, it is necessary to select or develop the instruments with which to obtain the information required to answer the questions.

The first and most difficult step is to decide what strategies are being implemented that discriminate the career education effort from other facets of the school program. By definition, an innovative career education program will possess characteristics that differentiate it from more traditional educational practices. These characteristics are the ingredients in the program's "recipe." If the "recipe" can be explicitly formulated, then it is likely that it can be replicated or improved. If it cannot, even if a program is judged successful in terms of the student outcome data, the staff will not know what dimensions of the program were responsible for success.

Specifying the treatments and relating them to participant groups was accomplished in the course of completing the tables and matrix discussed in chapter IV. Formulating the evaluation questions based on the tables and matrix was discussed in chapter V. The preceeding chapter dealt with obtaining measures of student

outcomes. This chapter discusses the identification of the sources from which treatment data may be obtained and approaches to securing that information.

A. Specifying Data Sources

Once the program treatments have been identified and appropriate questions posed, the next step is to identify the sources from which data will be collected to answer these questions. In specifying treatments and formulating questions, inevitably some consideration was given to the sources from which the answers were to come. In fact, if the questions were formulated precisely, more precisely than is often the case, this step will have been accomplished as an integral part of formulating the question. It is included here simply to insure that explicit consideration is given to alternative sources which may be possible and to assist in the selection or development of measurement instruments.

Generally, there are three possible sources of data for any given treatment question. Data may be obtained singly or in combination from:

- the provider of the treatment;
- the recipient of the treatment; and
- an observer of the treatment.

The first task in determining appropriate data source(s) for answering a given treatment question is to prepare a simple list of the groups of individuals in each of the above three categories. For example, a question might be: How does the classroom experience of participating third graders differ from non-participating third graders along the dimensions of (a) the amount of class time devoted to discussion of various jobs and their requirements, (b) the number of field trips to see workers in their work setting,

and (c) the number of class visitors who have discussed their jobs? Given this question, possible sources of the data necessary could be ordered as follows:

Providers: third grade teachers, school staff members who schedule field trips, and class visitors.

Recipients: third grade students.

Observers: school principal, Career Education coordinator.

Once a list of possible sources has been made, those sources from whom it would be feasible to collect data should be identified. The next task is to determine from which source the information is likely to be most accurate.

Once a decision regarding the data source has been made, it is necessary to identify the characteristics, if any, which differentiate the sources into important sub-groups. For example, it might be judged that the differences among teachers with respect to sex, age, number of training sessions attended, or attitude toward career education, might affect their responses to treatment questions. It may be of interest, for example, to relate selected teacher characteristics to the extent to which they utilize outside speakers or provide other treatments. Gathering this descriptive data regarding respondents will require supplementing the instruments to be selected or developed in the instrumentation step of the evaluation process with appropriate items.

The final task in identifying appropriate data sources is to determine whether the method of collection should be direct or indirect. In other words, must the individuals who will supply the data be asked in some manner for the information or does a system already exist which provides suitable information. If such a reporting system does exist, then the instrument selection/development step will consist of developing forms for recording data already available. If such a system does not exist, then the instrument task becomes somewhat more difficult.

B. Instrumentation

Once the treatments are identified, the questions formulated, and the data sources identified, the next step is to select or design appropriate instrumentation for collection of the data. Very rarely will the needed treatment instrumentation be commercially available or available from other career education programs because of the specificity and uniqueness of most treatments and program contexts.

While instrumentation typically must be site specific, it is possible to identify various types or approaches to instrumentation. Among the possible types of instruments which have been used in collecting treatment data are:

- document review summary forms;
- questionnaires;
- observation scales;
- staff reporting forms; and
- interview schedules.

Document Review Summary Forms: Data pertaining to some treatment questions may be included in documents regularly prepared by the school district for other purposes. For example, in some school systems it may be possible to assess the extent to which a career information center is utilized by various groups of students by reviewing the records kept by the center staff or school librarian. If the center has been in operation for several years and if comparable records exist, it may be possible not only to assess the extent to which particular groups have used materials during the evaluation period, but also to assess the extent to which they had used materials in prior years. By the same token, it may be possible to obtain information on the number and type of field trips taken in some schools by a review of school transportation records (if walking trips are not a feasible option in that situation). In many situations counselor records may also be a potent source of information.

As is true in development of all instruments, the key in developing document reporting forms is to know precisely what information is desired. Presumably this will have been accomplished in the process of identifying treatments and formulating evaluation questions. In addition to having a clear understanding of the information desired, it is wise to have a thorough knowledge of the documents which are to be reviewed prior to developing the forms. School files and other such documents are rarely designed to provide precisely the information desired by a program evaluation. As a result, it is necessary to develop summary forms which can obtain the maximum information available with the least amount of effort.

While the feasibility of obtaining the needed data from particular documents should have been a consideration in identifying the data sources, it may be necessary to reconsider the appropriateness of the decision once the form development process has begun. It is not unusual for a cursory review of documents to suggest that they contain information which upon closer analysis is found to be internally inconsistent or incomplete. It is also not unusual to discover that records do not report what they claim to be reporting. Frequently analysis shows that what appears on the records to be the number of students placed on jobs is in reality the number of referrals of students to potential employment. Another example of records not reporting what it seems often occurs when seeking the number of students receiving two or more treatments. The records may appear to indicate that 200 students are involved in various program activities when in fact a group of 50 students is receiving four different treatments. Further, it is not unusual for records to contain estimated figures which for most purposes are quite reasonable. However, for purposes of evaluation such estimates may not reflect subtle but significant changes and so may prove to be misleading.

In summary, while obtaining information from existing documents may at first seem to be the preferred approach on the grounds of ease of instrument construction and convenience to the school district, experience suggests that this is not always the case. Before concluding that this approach will be followed and finalizing the summary collection forms, it is necessary to go through a trial run. It may be possible that use of certain existing documents (e.g., counselor records) will not be possible or that which appears

to be an ideal format from the viewpoint of later analysis may be unduly cumbersome from the standpoint of searching the documents and recording the data. In all cases it is helpful, and frequently enlightening, to determine who completes a given record, when, and as a result of what action.

Questionnaires: Since all necessary information will be found in pre-existing documents, it is virtually inevitable that to obtain some information questionnaires will need to be developed. It is likely that separate questionnaires will be developed for teachers, counselors, students, administrative staff, and others who provide or observe treatments. As used here, a questionnaire is defined as an instrument which is given to an individual with the understanding that the recipient will carefully answer in writing the questions asked.

Ideally, a questionnaire would go through the same rigorous process of development as a test. In practice, however, generally this is not possible. Nevertheless, in the context of evaluating a career education program in a school district, the questionnaire will often be the most feasible method of accumulating answers to specific questions or sets of related questions from treatment providers, recipients, and observers. Below some guidelines for questionnaire development are provided.

In developing a questionnaire there are several basic rules which should be followed carefully. While they sound simple and have been repeated in the literature for years, they are violated frequently enough to warrent repetition. The primary rules are as follows:

- The questions should be simple and precise. Perhaps the most common difficulty in question construction is that what should be two or more

questions are combined. For example, if program personnel wish to ascertain how many of the students in a ninth grade English class that went on a particular field trip wrote an essay about it, several short questions are necessary. It is not enough to ask: "Did you write an essay on your field trip?" The first question should be: "Did you go on a field trip to X?"

- The questions should be understandable. The words and concepts should be readily understood by the reader. Ambiguous phrases should not be used. For example, if the need is to know how much time a counselor spent in career counseling, it is crucial to be certain that all recipients will use the same definition of the concept of "career counseling." If there is doubt, the question should be broken into various components.
- The criteria for providing the response should be clear. At times there are two or more perfectly logical ways to answer a question. For example, if the need is to know how many students were seen by a counselor once and how many were seen twice, it should be clear whether the number seen once is to be included in the number seen twice or whether the two groups should be reported separately.
- The response should require the fewest words from the respondent possible. Ideally, questions and possible answers can be so well thought out that the respondent will simply need to place a check mark in a designated square. This reduces responding and interpretation time. Often, however, this is not possible and open-ended questions are necessary. Where this is the case, tentative categories of responses should be prepared. In some cases it may be concluded in advance that there will be so little commonality in responses that an item should be excluded.

- Catch-all response items such as "other" should be avoided where possible. The existence of such items suggests that the question was not sufficiently thought out or that the information desired was not clear. It is not unusual for such items to produce no helpful information. There is also the danger that significant data will be lost because of the unspecific response.

In essence most questionnaires are collections of single questions put together such that they are all asked on the same form. The construction of the items and the order in which they appear then are the two most important factors. In item construction the preceding rules should be carefully followed. It may be useful to review questionnaires from other projects to identify items which they found successful. In using items developed elsewhere, consideration should, however, be given to the unique characteristics of the local respondents. The interpretation given to certain words or phrases may vary from location to location.

The primary consideration in ordering the items or sets of items are (a) their influence on responses to subsequent questions, and (b) the attitude engendered in the respondents. In the first instance, care must be given to the possibility that the answer to one question will dictate somehow the response to other questions when this is not desired. In the second, care must be taken to insure that the time necessary to respond to the first item will not lead the respondent to discontinue completing the questionnaire.

Observation Scales: In some situations it may be appropriate to develop instruments to be used by observers of classrooms or other project activities. The advantage of such an approach is that it permits an objective outsider to observe administration of the treatment; for example, the behavior of the teacher and pupils and their interaction. These observations may verify teacher reporting forms and/or may permit identification of facets of the classroom treatment which are not acquired through other data collection techniques.

The difficulties with this approach are two-fold. First, it is never possible to obtain as many observations as might be desired. Implementation is time consuming in that a third party must observe the treatment at least once and ideally several times. Second, if more than one observer is used, it is difficult, if not impossible, to be certain that all are using the same criteria in making their judgments. Observation of an event is subject to influences beyond the control of the observer and descriptions of the same event from the same person may change over time.

If this approach is utilized, it is important to develop observation scales which require as little subjective judgment as possible. As with questionnaires, the more often the observer can respond by the use of a check mark the better. Second, it is important, if several observers are used, that an assessment of inter-rater reliability be made. This simply means that several observers should view the same event and record their findings. The extent to which their reports are the same is the extent to which reliability among raters has been achieved.

Staff Reporting Forms: The purpose of a staff reporting form is to elicit from the direct provider of a treatment (e.g., the teacher), or indirect provider (e.g., coordinator), a description of relevant treatments on a continuous basis. Two distinct uses can be made of the results of this reporting process. First, the data provided can be used by administrators to monitor the activities of the program and to take corrective action as necessary. For example, if the number of field trips taken by teachers was reported monthly and if by Christmas some teachers had taken their classes on no trips, the career education coordinator could try to intervene.

The second use of a reporting form is to provide a summary record at the end of the year of the treatments which were administered to various groups of students. While this information might be secured through an interview or questionnaire given to staff at the end of the year, it is likely that the information will be more precise and accurate if the staff is reporting these events on an ongoing basis.

Given that a purpose of the evaluation is to assist local efforts improve, the two uses of the forms are quite compatible. As long as a record is made of the results of the interventions made by coordinating staff during the course of the year, using the forms for monitoring purposes will not affect the ability to use them at the end of the year in relating treatments to outcomes.

Given the advantages to program management that utilization of this approach provides, it is recommended for use to the extent feasible for any project. There are certain constraints on extensive use, however, and it should be used as a supplementary data collection device.

The difficulties in adopting the staff reporting form strategy extensively are of three types. First there is the problem of precisely identifying the information which is to be reported and the development of forms which capture the information in a consistent manner across all staff members. Since this is the same problem confronting developers of all types of instruments, it is not a sufficient reason to abandon this approach to data collection.

Second, there is the problem of accurate reporting of information. Since staff presumably will know that a use of the form is to monitor their activities, they may be more prone to provide less accurate information than on other instruments perceived to have less immediate impact. In part, this can be resolved through providing the staff with an understanding that what is sought is accurate information and that punative judgments will not be based on the information provided. In addition, simple procedures can be developed to verify the reporting information. These might rely on comparing the self reported information with school records indicating some of the same areas. For example, field trips reported could periodically be cross checked against school transportation records. Class observation is another method which could provide a periodic cross check on the results

of the self-reporting forms. Finally, students or others involved in the process could periodically be asked questions which would provide verification.

The third and most significant constraint is the imposition such systems make on staff time. If the system requires considerable time for the staff member to complete, there may be political or accuracy costs to pay. At the one extreme, teachers may simply refuse to complete the forms and at the other they may complete them carelessly and inaccurately. It is important, therefore, to develop forms which are quickly and easily completed and for which the staff can see a useful purpose. A number of career education projects have had success in utilizing this method in collecting a major portion of the treatment data and have developed forms which could be of use.

In developing staff reporting forms or procedures, most, if not all, staff objections and inherent difficulties mentioned above can be overcome if the reporting form is a natural by-product of normal action instead of something imposed in addition to normal procedures. If, for example, counselors write a simple referral sheet when referring students to jobs, it is better to have a carbon copy of that referral serve as the reporting form than to have a separate summary form for the counselor to fill out. Summation then becomes a routine task which can be performed by support staff on a regular basis.

Interview Schedules: The interview schedule or interview guide is a method of insuring that information which is best obtained orally is sought and recorded in a standard format. The interview schedule may be used to supplement any of the other methods of data collection either to verify the information reported or to add to it. For some purposes it may also be the only source of information.

Such schedules are either administered directly in a face to face interview situation or indirectly by telephone. In the first case, the format may either be quite tight, permitting little or no deviation from what is essentially a script for the interviewer, or may be quite loose permitting a conversational atmosphere, with a high probability that the precise order of the questions to be asked will not be as they appear on the guide.

The appropriate format to be adopted will depend on the use to which the information is to be put and the number and status of the interviewees. If there is a single interviewee (e.g., the school superintendent), then both methodological and pragmatic reasons suggest the looser format. The answers will not be directly compared with those of a similar individual and the superintendent is likely to resist a rigid interview format. On the other hand, if the purpose is to elicit standard responses from a group of teachers or students, then a more structured approach may be necessary.

As the sole source of a type of treatment data, there are three situations where this approach to data collection may be preferred. First, when the respondents are unlikely to respond to any other method, this obviously is the preferable approach. This may occur when information from employers providing training to high school students is necessary. Often there is greater possibility of obtaining a brief interview with an employer who supervises a student in a cooperative education program than there is of securing a written response to a questionnaire.

Second, when the respondents are expected to have difficulty with a written form based on linguistic difficulties. This may be due to their fluency in written English or the inherent ambiguity of some terms associated with the educational program. In both cases it may be judged essential that an interviewer have the opportunity to be certain that the respondent fully understands the questions asked.

Third, when the purpose is to explore an area with the respondent. In some cases it may not be possible to be precise with regard to questions in some areas. At times, what is sought is the outcome of following a chain starting from a precisely formulated question which could be answered in a multiplicity of ways. It may be, for example, that when the list of treatment questions was developed it was impossible to define precisely what happened in some locations. For example, it may not be possible to go beyond the general level of "distributive education -- grocery sales" at the start of the design process. In such a case it may be necessary to interview the teacher,

several students, and employers with regard to identifying the specific nature of the treatment. From the results of these open-ended interviews would emerge the needed description of the finer points of the treatment category.

Regardless of the approach or combination of approaches used to obtain information regarding a specific research question, it is clear that instrument development requires careful, time consuming labor. In all cases it requires precision with regard to the information sought and an understanding of the uses to which the information is to be put. It also requires an assessment of the practicalities of data collection such as ease of administration, and the time required. Most important, it requires attention to the issue of obtaining valid and reliable responses; the various cautions, guidelines, and suggestions contained in this chapter are designed to help in this regard.

Another factor to be considered is the ease of quantifying the treatment variables. Some would be relatively simple, such as the number of field trips or the number of different occupations represented by the outside visitors to a classroom. Other potentially influential factors, such as the emphasis placed by the teacher on career decision-making, may lend themselves only to very coarse quantitative categories, e.g., "much emphasis," "little emphasis," or "none at all." In the latter instances, it is necessary to have more than one judge assign these "quantitative" labels to the specific data, and to utilize the pooled judgments as the measure. Regardless of the method used for quantification, or the precision of measurement which may be possible, it is essential that each variable be quantified in order to be able to assess the relationship

between the treatment variable and its relevant outcome measure(s). Such assessment is only possible if outcome data can be grouped into at least two categories pertaining to treatment, such as "treated" and "not treated," or "high exposure to treatment" and "low exposure to treatment."

CHAPTER VIII
SAMPLING PLAN

The next step in the process is to specify the sampling plan. In an educational setting there are often many restrictions on how samples can be selected; this in turn will have important implications for the data analysis techniques which can or cannot be utilized (see Data Analysis Plan). One of the principal sampling restrictions is that it is often difficult to select a random number of students from a school or from pre-selected classrooms within a school. Often it is necessary to randomly select schools or intact classrooms within schools. If the evaluator decides to randomly select classrooms rather than students, then his unit of evaluation becomes the classroom, rather than the student. If four classrooms of 25 students each are selected, the size of the sample will be four, rather than 100. This fact is often overlooked by researchers who use the number of students as the sample size despite the fact that the logic of their statistical tests may require that the unit in the sample be selected on a random basis. A more complete discussion of the appropriate experimental unit in statistical analysis can be found in Julian Stanley's (ed.) Improving Experimental Design and Statistical Analysis. Our purpose here is to point out that the unit of analysis will vary depending on selection considerations.

Assuming that an adequate sample can be drawn, considering the design and unit of analysis, we may proceed to the next and most frequently asked question: "How big a sample must I take?" Since we always obtain more accuracy and more power in a test by increased sample size, a larger sample size would be better than a smaller one. In other words, with all other factors held constant, the larger the sample, the smaller the sampling error. The underlying question is "How much of a sampling error are we willing to put up with?" The answer depends upon: a) the consequences

of reaching wrong conclusions; and b) the cost of reducing the likelihood of being wrong.

The decisions involved in answering the above questions cannot be made on the basis of statistical or methodological considerations alone, since assessment of the consequences of error of given magnitudes and cost factors are essentially administrative considerations. Further, they depend upon the distribution of the variables in the population, and this is generally not known. As a reasonable approach to these problems, most studies of this type develop sampling plans which call for samples of approximately 5% (with a minimum of 30 if the group is relatively small) where the population numbers up to 5,000. For larger populations, over 5,000, a sample size of approximately 2% is generally considered adequate. In other words, sample sizes of between 30 and 250 where the population ranges up to 5,000 and of at least 100 for populations in excess of 5,000 are generally viewed as being stable enough to insure a sufficiently low error of estimate with a reasonable expenditure of time and money. Of course, these sample sizes may vary in accordance with the specifics of the research design. Either the 90% or the 95% confidence level is utilized to judge the significance of the findings.

The preceding rules of thumb with regard to sample size apply to each "homogeneous" group, e.g., students in the 9th grade who comprise treatment group X, or students in the 9th grade who constitute a "no treatment" group. It also refers to samples in which the students in the sample have been randomly selected. Where intact classrooms have been selected, i.e., where classrooms rather than students constitute the unit of analysis, it is necessary to select as many classrooms as possible to maximize the stability of the findings and to maximize the sensitivity of the statistical techniques which are applicable to small samples, i.e., samples between 10 and 20.

Sampling, then, requires a knowledge and implementation of procedures which allow for an unbiased selection of a small group that is characteristic of the larger

population such that estimates based on this smaller group can be generalized to the larger group from which it has been drawn. Of practical concern is the capability of drawing the sample according to the research design, the size of the sample, and an understanding of the statistical techniques that can or must be employed. It is not unlikely that the evaluator will find himself in a position of not having enough subjects, or resources, to investigate the large number of dimensions to be studied simultaneously. In this case, he may modify his design to accommodate the resources available. Matrix sampling, for example, might be explored in situations where there are relatively large numbers of students, the number of instruments (or items) to be used is quite large, and the time for testing is minimal. *

In summation, program personnel considering sampling issues should keep the following points in mind:

- Sampling is not an issue when all of the treatment population is tested.
- A census is the most accurate measure. Samples are taken when a census is not possible or practical. The larger the sample the better the measure.
- When the number to be tested is less than a census, random selection is necessary if the results are to be generalized.
- Confidence levels and therefore sample sizes can vary based upon practical or administrative considerations.
- As a rule, for treatment populations of 5,000 or under a sample size of 5% or 30 participants, whichever is the larger, is considered an

*Note: In general, the design of the sample is a very sophisticated problem requiring the expertise of a statistician. If no member of your staff has sufficient background in this area, an outside statistician should be engaged as an advisor.

appropriate or safe sample.

- The 30 or 5% rule applies for each treatment group.
- If classrooms are selected as the sample unit, it is likely that more students will have to be tested to insure a safe sample than if the student is the sample unit.

CHAPTER IX

DATA ANALYSIS PLAN

The specification of an analysis plan usually refers to the selection of the statistical method to be employed in analysing the data collected. Some statistical methods require that different kinds of data be collected than do others. For this reason, if the statistical method to be employed during analysis is specified prior to the collection of the data, it is necessary to consider the type of data required by that method prior to development of the sampling plan. In other words, the sampling plan would need to be designed so as to fulfill the requirements of the particular method chosen. As a practical matter the specification of the method to be used prior to development of the sampling plan serves primarily as a guide in developing the sampling plan.

Usually, however, the sampling plan will be developed around other considerations, such as time, money, practicality and availability of data processing resources. Most programs cannot afford the luxury of designing around analysis technique prerequisites. Therefore, the selection of the specific statistical methods to be employed is frequently based upon findings made after the data collection has been completed.

At this point in the evaluation process what is recommended is: (a) to specify the data processing requirements implicit in decisions regarding instrumentation and sampling, and (b) to prepare for the practical implications of the probable final analysis plan. In essence, this simply means to:

- estimate the time, personnel and equipment necessary; and
- arrange for its availability.

The first stage of data analysis begins immediately upon the completion of the collection of data. It involves making frequency counts and computing other descriptive statistics on all variables for each treatment group. Depending on instrumentation, getting to this point may or may not be a time consuming endeavor. This process should make possible the identification of non-productive variables, that is, variables with little or no response range. Since these cannot be of value in the analysis, they should be eliminated from further consideration.

The material which follows provides an overview of what statistical techniques are appropriate under what conditions. This is intended to serve as general background information for the next step of the analysis plan which is specifying the statistical techniques to be utilized.

Specifying the Data Analysis Techniques

Every statistical procedure for analyzing data is designed to answer a specific type of analytic question and is based upon certain assumptions regarding the nature of the data and the study design. The use of procedures designed for answering a different type of question than those discussed in these guidelines would be wasteful, while the use of procedures having data and/or design requirements which cannot be met would be erroneous and misleading. For these reasons, this section will focus only on those procedures which are apt to be appropriate; further, since the evaluation is oriented chiefly toward local use of the findings, it will focus on those procedures which yield findings that are relatively easy to interpret. This focus, however, is not intended to imply that expert statistical consultation may be dispensed with, whether at the design, data analysis, or data interpretation phase of the study.

Statistical techniques may be grouped into two large categories: parametric and non-parametric. The parametric techniques, which are more frequently used than the non-parametric, are based upon different assumptions, the most important of which is that the sample or samples must have been randomly drawn from one or more normally distributed populations. The advantage of these techniques is that they are more "sensitive" in detecting statistically significant findings (this is called the "power" of a test in technical terminology), providing that the sample is sufficiently large. For

small samples, and for larger samples which are not drawn randomly from normally distributed populations, it is necessary to utilize the less sensitive non-parametric techniques. By way of example, it is anticipated that the non-parametric will normally be used when dealing with data from the elementary grades, where classrooms, rather than students, are apt to be the unit of analysis. In this instance, the number of classrooms is likely to be small (e.g., 10 or 20) and the assumption of a normal distribution of classroom means on the outcome measures would be difficult to defend.

To answer the specific study questions formulated as Step 3 in the evaluation process, it is necessary to determine: (a) the magnitude of the pre-post change in test scores for each treatment group of students or classrooms, and (b) to assess the significance of the differences in change among the several treatment groups. This process is shown schematically in the diagram below:

	Scores on Pre-Test	Scores on Post-Test	Difference
Treatment Group	A	B	B-A
Comparison Group	C	D	D-C

Where the sampling plan justifies the use of parametric techniques, and where there are only two groups, the statistical technique known as the t-test for correlated measures would be applicable. If more than two groups are involved, and it can be assumed that students have been assigned to groups on a random basis, the basic analysis would be carried out by means of the analysis of covariance technique. Information concerning the pre-test

scores alone or the post-test scores alone would be obtained by the means of the analysis of variance technique. Where non-parametric techniques must be used, the Mann-Whitney U Test and MacNemar's Test for Change would be applicable if two groups are involved, and the Krushal-Wallis One-Way Analysis of Variance and the Chi-Square Test may be utilized to assess the significance of the inter-group differences.

Either within groups or across groups, correlational methods may be utilized to determine the "strength" of the relationship between treatment variables and outcome variables. For relatively large samples, the Pearson product-moment correlation would be applicable if there is a fairly large variation in the treatment variable; for small samples, the Spearman Rank Order Correlation would be applicable.

The potentially applicable techniques are described briefly below, and references to statistics books are cited for the interest of those who wish to investigate the technical aspects further.* The descriptions are provided first for parametric techniques and then for the non-parametric. They are arranged and numbered so that corresponding numbers indicate equivalence with regard to the type of study question for which the procedures are applicable. For example, the t-test (parametric technique 1) is equivalent to the non-parametric "Mann-Whitney U" (non-parametric technique 1).

---

* The sections on specific statistical techniques and the list of references is quoted, with permission, from Stenner, A. J., and Webster, W. J., Educational Program Audit., Arlington, Va., IBEX, Inc., 1971.

A. Parametric Techniques

1. The t-test

"The t-test is a statistical model that is designed to investigate several types of questions. The one-sample t-test determines, within specified levels of probability, whether or not the population from which the sample was drawn has a given mean. The two-sample t-test determines, within specified limits of probability, whether or not the populations from which the samples were drawn have the same mean. Different models are used, depending upon the type of sample data under analysis. In yielding a probability statement of the differences between group means, the model considers mean differences, sample variability, sample size, and whether the data are correlated or independent. (Blommers and Lindquist, 1960, chap. 12; Edwards, 1957, chaps. 13 and 14; Games and Klare, 1967, chaps. 11 and 12; Hays, 1963, chap. 10; Kendally and Stuart, 1967, chap. 11; Popham, 1967, chaps. 10 and 11; Wert, Neidt, and Ahmann, 1954, chap. 8.)"

2. The Analysis of Variance

"When two or more groups or samples are available, the analysis of variance is a model used to test for differences, within specified limits of probability, between the means

of those groups. The procedure enables one to analyze variances in such a manner that conclusions can be drawn about means. In the one-way situation (only one independent variable), the procedure is merely an extension of the t-test to situations where there are more than two groups. Higher order analysis of variance is used to ascertain, within specified limits of probability, the effects of two or more independent variables on a dependent variable. Utilizing higher order analysis of variance, the researcher can test main effects and interaction effects. In higher order analysis of variance, the researcher must concern himself with whether he is dealing with a fixed, mixed, or random model, since the type of model used will, in the presence of a significant interaction effect, make a difference in the error term used. Among the types of designs available with this technique are factorial designs, repeated measures designs, Latin Square designs, and numerous permutations and combinations of the various techniques. (Cochran and Cos, 1957; Cox, 1958; Edwards, 1957, chaps. 10 and 17; Edwards, 1960, chaps 9 through 13; Guenther, 1964; Hays, 1963, chaps. 12 through 14; Kendall and Stuart, chaps. 35 through 38; Lindquist 1953, Scheffe, 1959, Walker and Lev, 1953, chap. 9; Wert, Neidt, and Ahmann, 1954, chaps. 10 through 12; Winer, 1962.)"

3. The Analysis of Covariance

"When two or more groups or samples are available, the analysis of covariance is a model used to test for differences, within specified levels of probability, among the means of those groups, after initial compensation for differences among groups with respect to one or more control variables. Analysis of covariance is not a method to adjust for lack of random sampling; it is rather a technique that can be used to increase the precision of one's experiment if the control variable(s) selected is highly linearly correlated with the dependent variable. Analysis of covariance can be used in simple-classification or multiple-classification forms and, when the assumptions are met, is a more powerful technique than analysis of variance because it will provide a reduced error term. (Edwards, 1960, chap. 16; Guenther, 1964, chap. 6; Wert, Neidt, Ahmann, 1954, chap. 18; Winer, 1962, chap. 11.)"

4. Correlation Analysis

"When the words correlational analysis appear in the Metric-Metric cells of the taxonomy, they refer to the use of the Pearson product-moment correlation coefficient. Basically, correlational analysis provides the researcher with procedures for quantifying the measured relationships between two or

more variables. The size of the product-moment correlation coefficient varies from -1.00 to +1.00, thus providing the evaluator with an estimate of the size and direction of a given relationship.

"If the evaluator is interested in studying the relationship between two variables he could use the product-moment correlation coefficient. If he is interested in the relationship between one variable and a combination of two or more other variables considered simultaneously, he could use a multiple product-moment correlation coefficient. Multiple correlation provides an index of the relationship between a single metric variable and a composite. It is often evident that a relationship of some interest may be explained, at least to some extent, on the basis of correlations with a third variable or composite. In such a case, the evaluator may wish to determine the degree to which the two variables of primary interest are related beyond the relation implied by correlations with a third variable. The partial correlation between the variables of primary interest provides a measure of their relationship independent of some other variate. (Blalock, 1960, chaps. 17 through 19; Blalock, 1964; Edwards, 1957, chaps. 8 through 10; Guilford, 1965, chaps. 6, 14 and 16; Kendall and Stuart, 1967, chaps. 16, 26 and 27; McNemar, 1962, chaps 10 through 12; Walker and Lev, 1953, chap. 11; Zetterberg, 1963.)"

5. Regression Analysis

"Regression analysis is usually presented in the educational and psychological literature as a model for making prediction on a given criterion from one or a set of predictors. The yield of the approach is a simple or multiple correlation coefficient and a regression equation composed of a set of weights that can be used to optimize prediction. (Bottenberg and Ward, 1963; Draper and Smith, 1966; DuBois, 1957; Johnston, 1960; Kelley, Beggs, McNeil, Eichelberger, and Lyon, 1969; Kendall and Stuart, 1967, chap. 28)."

6. Discriminant Analysis

"Discriminant analysis represents an extension of regression analysis to the case where the criterion variable is descrete rather than continuous. The model is generally used for classification purposes. In essence, it provides an estimate of the position of an individual, based on specified information on that individual, on a line that best separates two or more classes or groups. Since it is often the case that one best line may not exhaust the power of a given group of measures for discriminating among groups, additional discriminant functions (to the lesser of the number of groups minus one, or the number of measures) may be fitted. Thus, the major purpose of discriminant analysis is to determine whether

discrimination among groups on the basis of a specified set of variables is possible or not, and then to reduce the size of the predictor space without substantial loss of information. Discrimination analysis itself does not define the regions of classification; however, approximate tests of the statistical significance of the separation of groups on a particular discriminate function are available, and the relative contributions of original variables to a discrimant function can be shown. (Anderson, 1958, chap. 6; Cooley and Lohnes, 1962, chap. 6; Kendall and Stuart, 1967, chap. 44; Morrison 1967, chap. 4; Rao, 1952, chap. 8.)"

B. Non-Parametric Techniques

1. Mann-Whitney U

"This non-parametric technique has been developed independently by a great number of writers to test, within specified levels of probability, whether the populations from which two independent samples have been drawn have the same distribution. It is one of the most powerful non-parametric techniques and is a useful alternative to the parametric t-test. (Bradley, 1960, chap. 5; Hays, 1963, chap. 18; Kendall and Stuart, 1967, chap. 31; Siegel, 1956, chap. 9; Walsh II, 1965, chap. 2.)"

2. Kruskal-Wallis One-Way Analysis of Variance

"The Kruskal-Wallis one-way analysis of variance is useful for determining, within specified limits of probability, whether or not the populations from which two or more independent samples were drawn share certain specified characteristics. It is an extremely powerful non-parametric technique and a useful alternative to the parametric one-way analysis of variance. Procedures for post-hoc analysis have been developed. (Bradley, 1960, chap. 12; Hays, 1963, chap. 18; Kendall and Stuart, 1967, 1967, chap. 31; Siegel, 1956, chap. 8; Walsh II, 1965, chap. 4.)"

(3. No equivalent technique)

4. Spearman Rank Order Correlation Coefficient (Rho)

"This statistic is a measure of association betwen ranks. It can be used for small or large samples, yields a coefficient between -1.00 and +1.00, and can be tested for statistical significance. It is an extremely powerful non-parametric technique for use with two ordinal variables. (Blalock, 1960, chap. 6; Hays, 1963, chap. 18; Kendall and Stuart, 1967, chap. 31; McNemar, 1955, chap. 12; Siegel, 1956, chap. 9; Walsh I, 1962, chap. 5.)"

(5. No equivalent technique)

(6. No equivalent technique)

7. Chi-Square

"This is an extremely flexible and often used non-parametric statistical technique. It can be used to test the quality of population proportions (Chi-square test for homogeneity) or for statistical independence among related populations (Chi-square test for independence). Chi-square can be applied to more than two populations and more than a dichotomization of the variable or variables of interest. This is the only practical non-parametric technique in which it is possible to have both more than two samples and data which fall into more than two categories in the same design. Post-hoc tests are available for both Chi-square for homogeneity and for independence. (Bradley, 1960, chap. 3; Hays, 1963, chap. 17; Kendall and Stuart, 1967, chap. 33; Siegel, 1956, chap. 8; Walsh II, 1965, chap. 5.)"

8. McNemar's Test for Change

"The McNemar's test for change determines, within specified levels of probability, whether or not observed changes in two related samples on the variable or variables of interest are greater than would be expected by chance. (Bradley, 1960, chap. 4; Siegel, 1956, chap. 4.)"

9. Sign Test

"The sign test is applicable to situations in which there are two related samples and the evaluator wishes to determine, within specified levels of probability, whether the populations from which the samples were drawn share certain specified characteristics. The test focuses on the direction of differences between each pair of scores and is generally used to attempt to detect a shift in location. Probabilities are determined by reference to the binomial distribution. The power-efficiency of the sign test is high for small samples (n=6) but declines as sample size increases. (Bradley, 1960, chap. 2; Hays, 1963, chap. 18, Kendall and Stuart, 1967, chap. 32; Siegel, 1956, chap. 9; Walsh I, 1962, chap. 7.)"

It should be stressed at this point that it is not the intention of this section to provide the reader with a "cookbook" set of instructions regarding data analysis. On the contrary, it is intended to make the point that statistical consultation is considered to be mandatory, both in formulating and in implementing the data analysis plan.

CHAPTER IX

REFERENCES

Anderson, T.W., An Introduction to Multivariate Statistical Analysis, New York: John Wiley & Sons, Inc., 1958.

Blalock, Hubert M., Jr., Casual Inferences in Nonexperimental Research. Chapel Hill, N.C.: University of North Carolina Press, 1964.

Blalock, Hubert M., Jr., Social Statistics, New York: McGraw-Hill Book Company, Inc., 1960.

Blommers, Paul and Lindquist, E.F., Elementary Statistical Methods. Boston: Houghton Mifflin Company, 1960.

Bottenberg, R.A. and Ward, Jose, Jr., Applied Multiple Linear Regression. Lackland A.F.B., Texas: Air Force Systems Command, PRL-TDR-63-6, 1963.

Bradley, James V., Distribution-Free Statistical Tests. WADD Technical Report 60-661, August 1960.

Cochran, William G. and Cox, Gertrude M., Experimental Designs, New York: John Wiley & Sons, Inc., 1957.

Cooley, William W, and Lohnes, Paul R., Procedures for the Behavioral Sciences. New York: John Wiley & Sons, Inc., 1962.

Cox, D.R., Planning of Experiments. New York: John Wiley & Sons, Inc. 1958

Draper, N.R. and Smith, H., Applied Regression Analysis. New York: John Wiley & Sons, Inc., 1966

DuBois, Phillip H., Mutlivariate Correlation Analysis. New York: Harper and Brothers, 1957.

Edwards, Allen L., Experimental Design in Psychological Research, New York: Holt, Rinehart and Winston, Inc., 1960.

Edwards, Allen L., Statistical Methods for the Behavioral Sciences, New York: Holt, Rinehart and Winston, Inc., 1957.

Games, Paul A. and Klare, George R., Elementary Statistics: Data Analysis for the Behavioral Sciences. New York: McGraw-Hill, 1967

Guenther, William C., Analysis of Variance. Englewood Cliffs, N.J.: Prentice Hall, 1964.

Guilford, J.P., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company, Inc., 1965.

Hays, William L., Statistics for Psychologists. New York: Holt, Rinehart and Winston, Inc., 1963.

Johnston, J., Econometric Methods. New York: McGraw-Hill, 1960.

Kelley, F.J., Beggs, D.L., McNeil, K.A., Eichelberger, T., and Lyon, J., Multiple Regression Approach. Carbondale, Ill.: Southern Illinois Press, 1969.

Kendall, M.G. and Stuart A., The Advanced Theory of Statistics, 3 Volumes. London: Charles Griffin, 1967.

Lindquist, E.F., Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.

McNemar, Quinn, Psychological Statistics. New York: John Wiley & Sons, Inc., 1962.

Morrison, Donald F., Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1967.

Popham, W. James, Educational Statistics. New York: Harper & Row, 1967.

Rao, C.R., Advanced Statistical Methods in Biometric Research. New York: John Wiley & Sons, Inc., 1962.

Scheffe, Henry, The Analysis of Variance. New York: John Wiley & Sons, Inc., 1959.

Siegel, Sideny, Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, Inc., 1956.

Walker, Helen M. and Lev, Joseph, Statistical Inference. New York: Holt, Rinehart and Winston, 1953

Walsh, James E. Handbook of Nonparametric Statistics, 3 Volumes New York: Holt, Rinehart and Winston, Inc., 1953.

Wert, James E., Neidt, Charles O., and Ahmann, J. Stanley, Statistical Methods in Educational and Psychological Research. New York: Appleton-Century-Crofts, Inc., 1954.

Winer, B.J., Statistical Principles In Experimental Design. New York: McGraw-Hill Book Company, Inc., 1962.

Zetterberg, Hans L., On Theory and Verification in Sociology. Totowa, N.J.: Bedminister Press, 1963.

CHAPTER X

DATA COLLECTION PLAN

The key to successful data collection is a well thought out plan that identifies the logistical activities in advance, provides a basis for logically selecting the personnel required to collect the data, and specifies time guides for judging progress. Table X-A illustrates a nine-step plan which can aid in designing a data collection effort that is appropriate for your Career Education evaluation study. Tables X-B, X-C, and X-D are provided at the end of this chapter as aids in developing the plan.

The first step in designing a data collection plan is to develop a list of all the tasks involved (the tasks may be listed in column 1 of Table X-B). In preparing the list, think of a task as anything that someone will have to do to carry out the data collection process. Do not overlook simple activities such as the reproduction of materials. Often these small tasks cause major problems for the entire evaluation effort. In recording all the tasks, list them in sequential order so that nothing will be overlooked. When finished, carefully examine the list to see that nothing is missing and that tasks flow logically from one to the other. Be sure that you are not expecting some activity to occur without having a task somewhere along the line to carry out that activity.

The second step is to establish completion or target dates for each of the tasks. (the completion dates may be entered in column 2 of Table X-B). These dates should be realistic and, if possible, provide some flexibility or cushion to allow for unexpected delays Be sure to take into account holidays, vacations, school activities, etc., when assigning completion dates to the tasks. With step two

completed you should know what has to be done and when.

Step three involves the identification of the skills needed to complete each of the tasks. (These may be recorded in column 2 of Table X-C.) One example of such a skill might be a knowledge of test administration. Another might be in the area of telephone interviewing.

The fourth step involves identification of the personnel available to carry out the data collection (These may be recorded in column 1 of Table X-D.) You will need to know the constraints on the use of various personnel. Some will have only certain hours or days available to work on the data collection. Another consideration will be the cost. Some personnel will have to be paid out of project funds. In these cases, the budget will have to be considered in deciding who to use. In addition, an assessment of the skills possessed by the available personnel should be made. (Constraints on the use of personnel may be recorded in column 2 and notation of their relevant skills in column 3 of Table X-D.)

The fifth step involves matching the personnel and tasks using the information you have developed in the previous steps. In some cases you may want to assign prime responsibility and also additional personnel for support activities (these assignments may be recorded in columns 3 and 4 of Table X-B).

The sixth step involves the determination of the training needs of the personnel you have decided to use. This can be done simply by examining

the required skills as you have outlined them in step three (Table X-C, column 2) and matching them to the skills of the personnel you have selected. Discrepancies between the two will have to be resolved by training the personnel to meet the project needs. (The discrepancies may be recorded in column 3, Table X-C.)

The seventh step requires the development of a training plan for providing the identified skills in the previous step.

The eighth step is the implementation of the training plan. It is not appropriate to include here detailed information on carrying out the training. However, the following example may be helpful. In cases where you are using instruments with which personnel are unfamiliar, there should be a training session to familiarize the staff with the instrument. An example of an outline for such a training session is:

A. Distribute samples of the test and answer sheets to the trainees.

B. Discuss the test, explaining all relevant aspects.

C. Give the trainees copies of blank answer sheets or use a device for projecting the information for all to see. Discuss the necessary identification items on the answer sheet and the methods by which the item responses are to be indicated.

D. Administer all or part of the test to the trainees.

E. Discuss the results and the interpretation of the instruments.

F. Clarify any questions the trainees have relative to the instrument.

G. Discuss the advantages and disadvantages of this particular instrument and any particular problem the trainee might have in administering the instrument.

Having completed the above eight steps, you have almost completed your data collection plan. The plan should be thought of as flexible and as a guide. When changes must occur, do not abandon the entire plan but change it to meet the new needs. A good data collection plan can be a real time saver but it almost always goes through a number of revisions.

As a last step you should evaluate the data collection plan after data collection has been completed and note any shortcomings that should be avoided during subsequent evaluation efforts.

In some instances all or part of the data collection will be completed by an outside contractor. When this occurs, steps 4 through 8 will be completed by the contractor in the proposal submitted to the program. In reviewing any such proposals the program staff should note the degree to which the points raised in these steps are considered and realistic.

When outside contractors are utilized the program will have to formalize the cost parameters. If no contractor is able to bid on the work to be done within the cost guidelines, either the work expected will have to be revised or the budget increased.

Table X-A

DATA COLLECTION PLAN

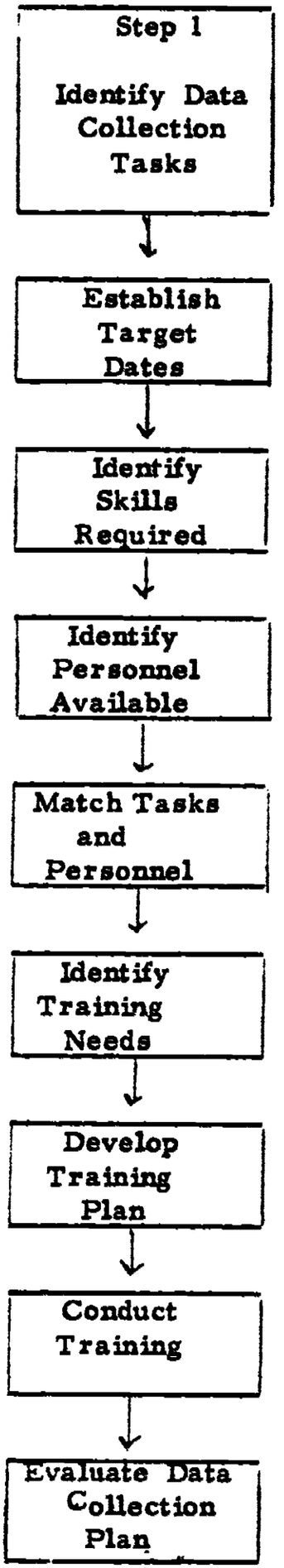


Table X-B

DATA COLLECTION
PLANNING CHART

(1) TASK	(2) Completion Date	(3) Prime Responsibility	(4) Support Responsibility

Table X-C

TRAINING NEEDS ASSESSMENT

(1)	(2)	(3)	
TASK	Skills Required	Skills Needed by Staff Member.	
		Skill	Staff

Table X-D

PERSONNEL PLANNING FORM

(1) NAME	(2) Restrictions - Cost, Availability, etc.	(3) SKILLS

CHAPTER XI
FINALIZE THE DESIGN AND PREPARE A REPORTING PLAN

Once the evaluation questions have been selected, instruments chosen or developed, analysis plans made, and data collection procedures established, it is necessary to review the evaluation design in its entirety in terms of costs and limitations. Once this is complete and any modifications made as a result, the final step in the design process is to prepare a plan for reporting the results of the evaluation.

A. Design Review

Throughout the process of specifying questions and instruments, and settling on appropriate sampling and analysis strategies, a series of decisions have been made which have possibly (a) redefined somewhat the original conception of the final products of the evaluation, and (b) affected the expected cost of conducting the evaluation. Before finalising the design, it is necessary to review this series of decisions to insure that the results are acceptable. Systematic recording of basic decisions on the Evaluation Design Worksheet presented at the conclusion of chapter II will greatly facilitate this process.

In most situations the final arbitor of design considerations will be the overall cost of the effort. While in theory it might be preferable to decide what needs to be done and then allocate the resources necessary to accomplish the task, in reality this is rarely, if ever, possible. Typically a ceiling cost of the evaluation was set well before the details of the design were specified and the issue at this point is to determine whether the ceiling was or was not exceeded. If it was not, then the review may focus entirely on substantive matters. If the absolute ceiling was exceeded, then compromises regarding

the substantive products of the evaluation must be made. This should be done at this point in the design process.

Once an assessment of the relationship between the cost of the evaluation to the resources available to conduct it have been made, the next step is to conduct a substantive review of the design. This review should begin with a reconsideration of the general questions which the various parties interested in the career education effort are likely to find of use and a determination of whether the design as constituted will provide the appropriate answers. It may be that for most people the design will yield direct answers while for others the information produced will be indirect or inferential at best.

In the likely event that certain questions are not directly answered, a determination must be made regarding the possibility, given the nature of the career education effort in the district and methodological considerations, of providing a direct response. In cases where this is not possible, then this should be stated explicitly. On the other hand, if a direct response was theoretically possible and a decision regarding instruments or sampling was made based on the time required to obtain the data or the availability or cost of instruments, then this too should be made explicit at this point.

If the assessment of required versus available cost of the evaluation indicates that more funds are available than are required to implement the design, then consideration should be given to refining the effort to answer the indirectly answered questions, or to addressing additional questions. On the other hand, if the cost requirements of the design exceed the available funds, then compromises must be made. Logically these will either be (a) to eliminate

altogether one or more questions, (b) to reduce the amount of effort expended on answering certain questions. For example, at this point it might be wise to reconsider the pros and cons of utilizing matrix sampling.

Once the review is complete, the assumptions of the design and its limits and strengths should be made explicit, i.e., a final version of the design worksheet prepared. This will provide the basis for assessing the extent to which implementation has been successful and for informing interested parties (e.g., the school board, teachers, etc.) what the expected results are to be. In addition, it provides the basis for completing the final step in the design process: preparing the reporting plan.

B. Reporting Plan

The data reporting plan will depend on those who will be using the evaluation report. It is often best to assume that a wide range of persons will want to read the report. Besides local project personnel and the U.S. Office of Education, the range of other persons might include:

- the local board of education;
- school administrators at all levels;
- teachers;
- guidance personnel;
- the state education agency;
- parents and other community residents; and
- other career education projects.

The report should be organized so that it can easily be interpreted by everyone who will want access to it. In many cases this may mean that special

summaries should be prepared for persons who do not need to know in detail the methodology behind the evaluation. A sample format is provided here.

I. OVERVIEW OF CAREER EDUCATION EFFORT

The basis for the overview should be the document developed in Task 1.

II EVALUATION QUESTIONS ADDRESSED AND SUMMARY OF ANSWERS

In this part of the report, the evaluation questions that the study addressed should be stated along with a summary of the findings that relate to the questions.

III. RECOMMENDATIONS FOR PROGRAM MODIFICATIONS

This part of the report should contain a statement or list of recommendations for program change or redirection. In some cases the evaluation may have raised other areas of interest for study. These areas of future interest should be addressed here.

IV. OUTCOME QUESTION/TREATMENT GROUP MATRIX

This part of the report should contain a copy of the matrix that was developed at the beginning of your evaluation along with a copy of the completed Treatment Group-Outcome Area Table.

V. DISCUSSION OF METHODOLOGY

This section of the report should include the evaluation design worksheets, copies of locally developed instruments, the data collection plan and noted deviations, and the final analysis plan. The discussion should include any deviation from the original design and any problems encountered in carrying out the design.

VI. DETAILED RESULTS

This section should include tables and figures displaying the data that were

collected for the evaluation.

APPENDIX

The appendix should contain copies of the instruments, if they are not included in VI above.

The first four sections of this report format can often be used as a summary document for distribution to those who do not need the detailed results and methodology.

In developing the outline for the report, attention should be given to actually constructing the tables that will be used. After these have been constructed, a review of the data tabulation forms should be made to assure that the information will be available for tabulation in the easiest possible manner. By examining these tables at this point, time can be saved in tabulation later on.

Depending on your evaluation design, it may be appropriate to provide interim reports. For instance, if information is being collected from teachers on a weekly basis regarding what they are doing in their classrooms, it may be helpful to report the findings of this activity at various intervals. There are two reasons for this. The information may be helpful to decision makers, and the teacher will respond better if he/she receives some feedback from his/her efforts. This is but one example of situations in which interim reports may be appropriate.

CHAPTER XII

IMPLEMENTING THE EVALUATION

A. Data Collection

Following the data collection plan outlined earlier is the key to sound data collection efforts. As a part of that plan, a planning chart was developed that will provide a task-by-task list of each activity to be carried out. The chart provides you with target or completion dates, and indicates the staff members who have prime responsibility and support responsibility for each of the tasks. The data collection plan also includes a training needs assessment and a training plan for the staff involved in the data collection. The training should be implemented at least a week before the first data collection takes place and alternative arrangements should be made for providing training to those who miss the regular training session. In some cases arrangements may have to be made to provide separate training for persons with different evaluation responsibilities who work a different schedule, or are physically separated from the rest of the evaluation personnel.

The types of data to be collected are basically two; these are:

- activities that occur only one time during the course of the evaluation such as pre-test, post-test, certain observations and interviews.
- activities that are continuous and usually spread over the life of the project. An example of one such activity is the regular reports that a teacher might submit as a part of the evaluation. These reports are repeated over and over on a routine basis.

For the second type of data collection, it is recommended that the evaluation staff routinely monitor the data that is being submitted. There are two reasons for the monitoring of this information. The first is to assure the quality of the data that is being collected. If the quality goes down, corrective measures can be taken that will provide a higher quality of data. The second reason is to give feedback to the persons collecting and submitting the data. The feedback is often helpful to the person providing the data and makes the person more likely to cooperate in all aspects of the evaluation.

Hopefully, all the tasks related to the collection of data will be completed on time and without any difficulty. However, in an evaluation process, various events can occur that are outside the influence of the evaluators. The evaluators must always try to anticipate these events and be flexible enough to overcome these obstacles. The following are some examples of typical problems that may occur:

- A flu epidemic the day of the tests.
- A lack of cooperation from participating staff or students.
- Personnel that have been closely involved in the evaluation are suddenly transferred.
- New program activities are added that impact on certain outcomes.
- Budgetary changes are necessary that will reduce evaluation monies.

Each of the preceding problems can have a serious effect on the evaluation study. Of course, there are only a certain number of precautions that one can take. For example, by having staff closely involved in the design of the evaluation, the chance that staff will not cooperate is minimized. One cannot, however, predict the flu. Handling these problems requires the best judgment of the staff. In some cases, one might have to modify all or part of the design. In making any change in the evaluation design, it is important to record accurately exactly what did occur so that there is a record to refer to during the analysis phase. This record of deviations from the original design should be included in the evaluation report so that those using the information contained in the report understand the context in which the findings are to be considered.

The last step in implementing the data collection plan should be an evaluation of the data collection plan and all data collection activities. Often, people involved in evaluation make the same errors year after year. Ideally, every person involved in the data collection should be involved in the evaluation. If a number of the "bugs" in the data collection plan become apparent on the day that the major collection activity is taking place, these should be noted promptly. Other "bugs" will turn up when the data are tabulated. An evaluation will note any deficiencies and pinpoint or suggest their causes.

B. Data Processing And Analysis

The processing needed to prepare the data for statistical analysis should be implemented as soon as possible after the first batch of data has been collected. This would include any necessary labeling for identification purposes, coding, and key-punching, if machine processing is to be used. The process should be undertaken early for several reasons: a) the system governing the flow of the data can be improved early, b) certain defects in the data can be corrected while the facts and circumstances are fresh in people's minds, and c) labeling or coding difficulties can be discovered early, minimizing the need to make corrections on large amounts of data which have already been processed. Insuring that labeling is error-free is of particular importance in a pre/post type of study because of the need to link or pair data from the post-test phase with data from the pre-test phase on a student-by-student and a class-by-class basis. This aspect of data processing should be carefully developed or reviewed by the computer programmer if machine analysis is to be used, or by the statistical consultant if the analysis will be done manually.

It should be anticipated that the realities of the data collection effort are likely to require modifications in the data analysis plans; these would stem chiefly from deviations from the original sampling plan and from the fact that incomplete data will have been obtained from various students or various classes. Techniques exist for handling or accommodating most of these deviations, and it is essential that the analysis specialist be kept informed of such deviations as they occur. Modifications in the data analysis plan can also be anticipated on

the basis of preliminary tabulation o early data; such tabulation would display basic facts about each measured variable, such as means, standard deviations, ranges, and sample sizes for each homogeneous grade and treatment group of students and classes. Other descriptive statistics may be useful for internal program use and/or for refining the data analysis plan; the procedures for obtaining these and all other tabulations would be developed by the researcher or the statistical consultant.

It should be pointed out that most studies in education leave insufficient time for carrying out the data analysis. The more time that can be allocated for analysis, the greater will be the opportunity to pursue any leads, hunches, or questions which each phase of the analysis might suggest. The plan, in other words, should be viewed as providing the basic framework for analysis, rather than as a rigid prescription. In this way, the data will yield the maximum degree and type of meaningful information and important cues for interpretation or explanation of the basic findings.

C. Reporting Findings

In the broadest sense, the purpose of evaluation is to provide information concerning performance to decision makers and planners preparing for the future. The evaluation process is not complete until the knowledge gained is available to be utilized by planners, managers and policy makers. This, of course, is as true for career education programs as it is for others. Because career education is a relatively new effort, there does not exist large amounts

program information. Therefore, the reporting and dissemination of evaluative information is important for individuals seeking to improve career education programming at all levels.

Different kinds of information are needed at different levels of planning and administration; the reporting plan should take this into account. As a rule, individuals at the operating level (teachers and counselors) require specific information on each student (individual data) while policy makers at the school board level need more general information (aggregate data). At the state and federal level the information need becomes increasingly general. Clearly policy makers and planners at higher levels do not need, nor could they handle, detailed and specific data from many sources. The need for increasingly general information at higher policy making levels, carries with it a need to make comparisons between programs (treatments), their settings (context) and students (participants). The evaluation and reporting plan has been designed to provide information which will be useful to all levels of the career education program.

By classifying local objectives and treatment in terms of Outcome Areas and broad activity areas it is possible for planners and policy makers at the federal level to compare types of programming and results without being faced with the impossible task of trying to make judgments based upon many individual case histories.

To assist those not familiar with the context in which the career education effort being evaluated occurred,the provision in the report of certain general, descriptive information would be helpful. Particularly valuable in this regard would be information regarding the size of the school district, and the socio-economic characteristics of the area.

In state or federally funded programs, the funding source will usually stipulate the kinds of information to be reported. In general, however, in reporting findings it is important to describe changes that have occurred in the target populations. It is also important to describe the relationships between planned change and those changes which actually occurred. In reporting change, it is necessary to describe what produced that change. In other words, what was different for one student or group of students than for others. Or, what was different this year as compared to last year. Without this descriptive information it will not be possible to understand the significance of any change that actually occurs. Information that relates to the difference between planned and actual change is important in understanding which treatments or approaches are most effective, etc. The value of this kind of information to planners at all levels should be clear.

Writing a report without disseminating that report is wasteful. You should consider disseminating your results to as many groups as possible. Even if your program is not funded by USOE, it would be worthwhile to send them a copy. You are strongly urged to send a copy of your final report to the ERIC Clearinghouse for Career Education (Northern Illinois University).

APPENDIX

A. Acknowledgments

Preparation of this draft plan for evaluating Career Education has involved many persons and the review of much literature. Development Associates staff has been assisted in this process by the members of a panel of career education program third party evaluators, members of a panel convened to review instruments measuring student outcomes, officials within the U.S. Office of Education, and the individual contributions, over and above their participation on the evaluators' panel of : A. Jackson Stenner of IBEX, Inc. and Jerry Walker of the Center for Vocational Education, Ohio State University.

The members of the panel of third party evaluators, held on July 18 and 19, 1974, at Development Associates' Washington Office were:

Carol Aslanian - Policy Studies in Education. New York

Mary Belcher - Educational Services, Inc. Waco, Texas

Jerome Moss - University of Minnisota, Minnisota. RCU Minneapolis, Minnisota

Lee Olson - Marshall University. Huntington, West Virginia

Hugh Peck - IBEX, Inc., Durham, North Carolina

A. Jackson Stenner - IBEX, Inc., Durham, North Carolina

Jerry Walker - The Center for Vocational Education, Ohio State University. Columbus, Ohio

The instrument review panel held on August 5 and 6, 1974, at Development Associates' Washington Offices considered 90 instruments which have been used to measure career education student outcomes. A quite promising set of instruments, The New Mexico Career Education Test Series published by MONITOR, was received after the panel had adjourned. Copies of these instruments were sent to members of the panel for their comment. Their comments are expected to be included in revisions to the draft. Also to be included in revisions will be answers as they are received from the instrument developers, to questions raised by the panel. Two such responses are included as sections C, and D below. The members of the review panel were:

Dr. Nancy Burton:

Scoring Director, Department of Research and Analysis, National Assessment of Education Progress, Educational Commission of the States. Formerly responsible for development of the National Assessment's effort in Career Occupational Development.

Dr. William E. Coffman:

Lindquist Professor of Education and Director of the Iowa Testing Programs at the University of Iowa. Member of the Analysis Advisory Committee, National Assessment of Educational Progress. President of the National Council on Measurement in Education (1972).

Dr. Edward Lareau:

Associate Director for Research, Research Coordinating Unit, Admiral Peary Area Vocational-Technical School, Ebensburg, Pennsylvania.

Director of the Career Education Project sponsored by the area vocational-technical school.

Dr. Dale Prediger:

Research and Development Division of the American College Testing Program. Directed the development of the Assessment of Career Development.

Dr. Donald Super:

Professor, Columbia Teachers College. Pioneered development of the concept of vocational maturity. Developer of the Career Development Inventory.

In addition to the above, input in several areas from career education practioners attending a series of 20 mini-conferences sponsored by the U.S. Office of Career Education was obtained.

B. Partial Bibliography

A large number of reports, monographs, and books in the area of evaluation of educational and other public programs was reviewed in the course of developing this draft. Aside from general texts cited in the body of the

draft,evaluation reports from over 50 career education projects, and evaluation guidelines for several federal programs, sources of particular note are listed below. The draft plan reflects a syntheses of a multitude of ideas contained in these citations.

Agency for International Development. Evaluation Handbook. Washington, D.C. February 1972.

Anderson, S.B., Ball, S., Murphy, R.T., & Rosethal, E.J. Anatomy of Evaluation: Important Concepts and Techniques in Evaluating Education/Training Programs. Educational Testing Service (Sponsored by the Office of Naval Research. Princeton, N.J. August, 1973.

Hatry, Winnie R., & Fish, D. Practical Program Evaluation for State and Local Government Officials. The Urban Institute. Washington, D.C. 1973.

Kosecoff, J. &Fitz-Gibbon, C. "Many A Slip", Evaluation Comment. Center for the Study of Evaluation. December 1973.

Peck,H. &Stenner, A.J. Instrumentation and Measurement Strategies for Information Based Evaluation (Working Draft). IBEX Inc., Arlington, Virginia. 1973.

Phi Delta Kappa Study Committee on Evaluation (Daniel Stufflebeam, et.al.). Educational Evaluation and Decision Making. F.E. Peacock Publishers, Inc. ltasca, Illinois. 1971.

Stenner, A.J., & Webster, W.J. Educational Program Audit Handbook. Available through: IBEX, Inc., Arlington, Viringia. 1971.

Stenner, A.J. An Overview of Information Based Evaluation: A Design Procedure. IBEX, Inc., Arlington, Virginia. 1972.

Stenner, A. J. Reporting Evaluation Information: Practical Alternatives and Guidelines (Working Draft). IBEX, Inc., Arlington, Virginia. 1973.

Suchman, Edward A. Evaluative Research: Principles and Practices in Public Service and Social Action Programs. Russell Sage Foundation. New York. 1967.

Weiss, Carol H. "Between the Cup and the Lip", Evaluation Comment. Center for the Study of Evaluation. December 1973.

C. Response Regarding the Self Observation Scales

The attached comments were received in response to questions raised by the instrument review panel from IBEX, Inc.

In conversation with National Testing Service, Inc. regarding scoring of the SOS, the writer was informed that (1) SOS is scored by the "complete factor" method with each item weighing on each scale; teachers or others would not be able to score the SOS, (2) Scoring and reporting services are included in the purchase cost of the instrument, (3) Any attempt to hand-score would invalidate the factor structure, (4) Complete information regarding these areas is available in a series of technical papers by the authors from National Testing Service, found elsewhere in this document.

The Achievement Motivation scale is experimental. Whereas the other scales were determined by factor analysis, Achievement Motivation was determined by discriminant function analysis. Work on the validation and replicability of Achievement Motivation is continuing.

D. Response Regarding the Differential Aptitude Test.

The attached table was provided by the Psychological Corporation to illustrate group summary data which they are prepared to make available to facilitate use of the Differential Aptitude Test in program evaluation.

Model for Group Summary of Data Available for the Differential Aptitude Test

In relation to first occupational choice, Educational Aspiration is:	Relation between first occupational choice and school subject preferences	Relation between first occupational choice and relevant aptitudes	Distribution of Cases	
			N	%
Low	No Match	Qualified	XXX	XX
High	No Match	Qualified	XXX	XX
Low	No Match	No Match	XXX	XX
High	No Match	No Match	XXX	XX
Low	Match	No Match	XXX	XX
High	Match	No Match	XXX	XX
Low	Match	Qualified	XXX	XX
High	Match	Qualified	XXX	XX
Low	Match	Over Qualified	XXX	XX
High	Match	Over Qualified	XXX	XX
Low	No Match	Over Qualified	XXX	XX
High	No Match	Over Qualified	XXX	XX
Match	Match	No Match	XXX	XX
Match	Match	Qualified	XXX	XX
Match	Match	Over Qualified	XXX	XX
Match	No Match	No Match	XXX	XX
Match	No Match	Qualified	XXX	XX
Match	No Match	Over Qualified	XXX	XX

Similar tables would be prepared for the Second and Third Occupational Choices.