DOCUMENT RESUME

ED 034 071
VT 009 795

AUTHOR
Spiess, Kathryn H., Ed.; Spiess, Eugene R., Ed.

TITLE
A Guide to Evaluation; Massachusetts Information Feedback System for Vocational Education. First Technical Progress Report.

INSTITUTION
Massachusetts Research Coordinating Unit for Vocational Education, Boston.

Pub Date
Sep 69

Note
121p.

EDRS Price
EDRS Price MF-$0.50 HC-$6.15

Descriptors
Accreditation (Institutions), Cost Effectiveness, Educational Objectives, Evaluation Criteria, Evaluation Methods, *Evaluation Techniques, *Information Systems, Models, *Program Evaluation, Questionnaires, *State Programs, *Vocational Education

Identifiers
*Massachusetts

Abstract
        A total educational information system for evaluation of vocational education in Massachusetts is described. Specifically, the evaluation guide describes the evaluation plan, reveals the philosophy of evaluation upon which the design was built, outlines the processes of evaluation called for in the design, introduces the forms for data collection that will be used in the design, and acts as a reference manual for those persons at the local level responsible for conducting specified segments of the design. The system is presented in four phases: (1) Program Evaluation, (2) Process-Product Evaluation, (3) Cost-Benefit Evaluation, and (4) Evaluation of Vocational-Technical Education (Impact Study). (CH)

A GUIDE TO EVALUATION

MASSACHUSETTS INFORMATION FEEDBACK SYSTEM FOR VOCATIONAL EDUCATION

INTRODUCTION

This publication is essentially a technical report describing the development of a total educational information system for vocational education in Massachusetts. A multi-audience on both the state and national level has been considered in drafting the publication, including: vocational educational administrators on the federal, state, and local levels; vocational education teachers; students of vocational education; research and evaluation specialists; and the concerned lay public. Our staff would be delighted to respond to any inquiries concerning developments relating to any phase of the total educational information system.

Both the Massachusetts Information Feedback System and this publication owe their existence to the genius and tireless efforts of Jim C. Fortune, University of Massachusetts.

William G. Conroy, Editor

TABLE OF CONTENTS

PREFACE

The evaluation of Vocational-Technical Education is a
primary concern of Congress and those charged with the respon-
sibility of administering Vocational-Technical Education, as
are other programs partially or fully supported by Federal
dollars. Public Law 90-576, Amendments to the Vocational
Education Act of 1963, states that the State Advisory Com-
mittee shall "evaluate vocational education programs, ser-
vices, and activities assisted under this title and publish
and distribute the results thereof: and prepare and submit
an annual evaluation report, accompanied by such additional
comments of the State Board as the State Board deems appro-
priate, which (1) evaluates the effectiveness of vocational
education programs, services and activities carried out in
the year under review in meeting the program objectives set
forth in the long-range program plan..."

During its November, 1968 conference the Vocational

Education Research Initiation Team of the Research Coordination Unit for the Commonwealth of Massachusetts determined that the development of an evaluation process for Vocational-Technical Education was the second most crucial issue (second only to the education of the disadvantaged) facing Vocational-Technical Education at this time in the state.

Acting under the auspices of the new amendment and the stated concern and need by the Research Initiation Team, the Research Coordination Unit of the Commonwealth of Massachusetts has attempted to design a state-wide evaluation that will (1) fulfill the criterion of providing a state-wide data base for the assessment of Vocational-Technical Education programming, (2) meet the needs of local institutions in terms of offering viable feedback upon their programs, (3) continue to grow and be flexible enough to meet the increasing needs for evaluation caused by program growth, (4) gather data on the three most essential aspects of Vocational-Technical Education, namely, product, process and cost, and (5) allow for decision-making at the local as well as at the state level. The evaluation guide seeks to: (1) describe the evaluation plan, (2) reveal the philosophy of evaluation upon which the design was built, (3) outline the processes of evaluation called for in the design, (4) introduce the forms for data collection that

will be used in the design, and (5) act as a reference manual for those persons at the local level responsible for conducting specified segments of the design.

CHAPTER ONE

INTRODUCTION TO EVALUATION

Jim C. Fortune

University of Massachusetts

Analysis of Current Practices

Evaluation in education has failed to develop as rapidly as other phases of the educational process because educators have only partially adopted scientific methodology. This failure to adopt scientific methodology is due in part to the nature of the variables with which educators must deal, in part to the types of people which enter into the educational profession, and in part to the mistrust and misconceptions existing in the field concerning empiricism and more specifically evaluation.

The evaluation of educational programs has often been confused with justification and the judgment component of evaluation has been viewed by educators as a threat. Even though the major charge of evaluation, according to current thinking, is to provide data needed for decision making, the

2

required evaluation of funded programs has led educators to confuse evaluation with refunding criteria. Hence, program administrators have felt it necessary to show goal attainment as a basis for refunding, making evaluation a _post hoc_ operation which is primarily concerned with searching for positive program effects after the program has been completed. This use of evaluation in an after-the-fact manner robs evaluation of its most potential benefits and places the evaluator in a role of having to pass judgment upon a program which cannot be altered from the information produced. In viewing evaluation as a _post hoc_ operation of which results can do little to alter the effectiveness of a program or to point out weaknesses of the program in time for them to be corrected, the program personnel are threatened by the finality of the results and must reject evaluation as invalid in order to protect their ego-investment. With, however, the absence of a commonly-agreed-upon criterion for "good" teaching and with the inferential nature of variables related to behavior change, the threatening aspects of value judgments based upon _post hoc_ evaluation become understandable. Such threats can easily be erased, however, by applying evaluation as a feedback mechanism rather than a _post hoc_ operation, allowing evaluation to offer directions for program modification and operation.

The structure of public education has also been the basis
for mistrust and misconceptions concerning evaluation. The
rigidity of the structure of the public education system, at
times, inhibits proper evaluation processes and makes eval-
uation difficult if not impossible. The public nature of
education has tended to put educators on the defensive and
has made them feel the need to overly justify their efforts
hindering the communication needed for proper evaluation.
The time demands and public pressures placed upon educators
have not allowed them to spend the necessary time required
to plan and carry out effective evaluation. Program commit-
ments and the inflexibility of program design leaves little
room during the process for effective modifications. Previous
practice of only paying lip-service to evaluation has left
the field of practitioners with remembrances of the ineffec-
tiveness of the evaluation process. The practice of justifying
programs through the testimonies of experts and through the
sanction of accreditation agencies have led educators to believe
that these activities fulfill the evaluation needs of a program
without ever questioning the kinds of effects the program has
on students or on staff.

Up to the present time, teachers, administrators and
counselors have received their training in evaluation techniques

on a piecemeal basis. Some of their training was received in courses in tests and measurements; some of their training was acquired in educational statistics and psychology courses; and other parts of their training were tacked on to methods courses. This piecemeal approach to the development of evaluation skills has generally been unsatisfactory. Educators of the field are normally poor evaluators, part of which is due to the inability to synthesize the varying fragments of training into functional skills. There are often few attempts to analyze what standardized tests measure. Very seldom is there any effort made to see whether or not the items on locally-made tests communicate clearly. Usually, evaluation procedures for instructional programs take place only at the end of the program, too late to become part of the decision-making feedback process which should be an integral part of evaluation. Too often single-measures possessing little resemblance to terminal behaviors (because paper and pencil tests seem to be the easiest route to criterion evaluation) are used and contextual variables are often ignored.

Confusion in regard to purposes and usefulness of evaluation has generated a lack of confidence in the ability to evaluate programs functionally. Heavy commitments of personnel limit the amount of time they have to evaluate their efforts

adequately. The difficulty of assigning causal relations to variables and of identifying cost factors of programs have left program decisions for revision and selection in the realm of mysticism.

Review of current thinking

Curriculum revisions, new designs in teacher education, and other innovative programs are evidence of the search to improve education. Inherent in these attempts is the assumption that the status of the education system is known and that its inadequacies can be empirically identified. These attempts to improve instruction and instructional processes, however, have served to spotlight evaluation weaknesses and indicate that most educational improvement programs are carried out on a trial and error basis. Without the ability to diagnose student capabilities and needs, without the ability to measure and describe program outcomes, without the ability to compare alternative programs, there exists little basis for directed improvement. In "Learning for Mastery" Bloom (1968) posits:

> There is little question that the schools now do provide experience for some students--perhaps as high as one-third of the students. If the schools are to provide successful and satisfying learning experiences for at least 90 per cent of the students, major changes must take place in the attitudes of students, teachers, and administrators; changes must also take place in teaching strategies and in the role of evaluation.

6

The needed changes in evaluation perceived by Bloom have been adequately pointed out and described by Guba (1968) and by Stufflebeam (1968). In order that evaluation serve better the needs of education, it must become a more active part of the process of change. Evaluation must become an important activity in decision-making. Feedback for curriculum revision should come from evaluation. Curriculum choices should be based upon data generated for evaluation. Hence, the evaluation needs of the education system are more than the limited tests and measurement skills usually offered. Evaluation must become more than the quickly gathered "after-the-fact" data selected to justify a program. Evaluation must become descriptive, active, and better utilized in the schools.

In the training of personnel to carry out evaluation in today's school programs, the emerging role of evaluation must be given more prominence. No longer can education rely on the indirect method of training evaluators through the development of measurement skills. In the early 1930's, tests and measurement were adequate as a beginning of educational evaluation. Lord, however, recognized that tests and measurements were only part of the skills needed in evaluation and that evaluation was a much broader concept. (Pace, 1968). Since

7

teachers and other educational personnel tend not to utilize the tests and measurement skills usually taught in education programs, evaluation skills with easily recognized relevance may be improved and become the route through which educational evaluation can become more useful.

Educators tend to be overburdened with knowledge to be gained and skills to be acquired. Perhaps, evaluation skills should be part of the curriculum for only those who wish or need them in their repertoire, even though we might wish to impart an awareness to all. Specialists and/or technicians may well be able to increase the efficiency of evaluation and to improve the quality and usefulness of evaluation in the improvement of education.

In the establishment of a curriculum and program in evaluation, care must be taken to include all of the skills needed to carry out the broad process called "evaluation". Cronbach (1962) defines "evaluation" as the collection and use of information to make decisions about an educational program. Stake (1967), in describing evaluation, says:

> Both description and judgment are essential—
> in fact, they are the two basic acts of evaluation.
> Any individual evaluator may attempt to refrain
> from judging or from collecting the judgments of
> others. Any individual evaluator may seek only to
> bring to light the worth of the program. But their
> evaluations are incomplete. To be fully under-
> stood, the educational program must be fully described
> and fully judged.

8

Stufflebeam (1968), working out of the Evaluation

Center at Ohio State University, emphasizes the role of

evaluation as a tool in decision-making. His analysis

breaks evaluation into four categories; context evalu-

ation, input evaluation, process evaluation, and product

evaluation. In the so-called CIPP model of evaluation,

context evaluation refers to the assessment of variables

that determine the school's status, that characterize the

school's setting and that are related to the identification

of needs and problems. Input evaluation refers to a

description of the resources and components available and

used in coping with the specified problems. Process evaluation

is defined as feedback mechanism used to determine the effec-

tiveness of the program and its individual components. In

process evaluation, the educational program of the instruc-

tional sequence is evaluated in terms of objective, purpose,

and/or a set of operating standards. The fourth breakdown

of evaluation is product evaluation, which includes the deter-

mination of the feasibility, quality, and efficiency of the

program in terms of its product, i.e. students.

Pace (1968) of the Research and Development Center on

Evaluation of UCLA prefers a categorization system of eval-

uation activities organized around the size of the program

9

being evaluated. He is careful to point out that evaluation
strategies and the skills needed to carry out the evaluation
activities differ significantly as size of unit, scope of
unit, and duration of the unit being evaluated vary. For
small units (such as programmed tests, instructional methods,
or instructional units) relevant evaluation can be directly
related to behaviorally defined objectives, largely limited
to intended effects, designed as a hypothesis testing exper-
iment, directly criterion based, and largely unconcerned with
group processes. For a medium sized unit (such as, a particular
curriculum, a single school, or a single grade) relevant
evaluation should not be limited to explicitly defined program
objectives, but include a range of potential outcomes and inter-
actions, sometimes approximate an experimental mode (controlled
conditions for the duration of the program are rarely possible)
and be concerned with group processes. For large units (such
as a school system, a total institutional program, or higher
education in the United States) relevant evaluation should never
be designed as an "experiment" and be concerned with group
processes only when relevant.

Following is a summary of the major emphases concerning
evaluation at this time:

1. There is a need for evaluation to become more
   descriptive of what is going on in the schools.
   This information as to what is must be fed back
   into the system to produce further change (which
   in turn requires further evaluation). We can no
   longer view evaluation as completed when the final
   report is written.

2. Tests and measurements no longer describe evaluation.
   There is need to train those with evaluation
   responsibility in new procedures of evaluation to
   improve the quality and usefulness of evaluation.

3. Evaluation is a tool in decision-making. It
   describes what is going on. The decision as to what
   steps to make after the evaluation are often human-
   istic, existential decisions; but, evaluation makes
   it possible to define the effectiveness of these
   decisions.

4. The broad movement of new thinking in educational
   research appears to be moving toward a process view
   of evaluation in which evaluation never ends, but
   is on-going. This process view is meaningful to
   the conduct of classrooms, special projects and the
   curriculum in general, and may provide an effective
   link for the first time between educational researchers
   and school personnel.

5. The procedures and techniques of evaluation will
   vary as the scope and size of the evaluation changes.
   The larger the system, the more a "process" approach
   to evaluation may be required.

Role of Evaluation - A Basis for Directed Change

As can be seen from the current thinking concerning

evaluation, there is a natural role in the instructional process

for evaluation. Evaluation serves as the basis for directed

change in education. Through evaluation educational programs

can be described adequately for replication, diagnosed for

strengths and weaknesses, validated for effectiveness, modified
and redesigned for efficiency, and judged for relevance and
pertinence.

C. Robert Pace in "Evaluation Perspectives: '68,"
outlines several components of evaluation that indicate the
total power of evaluation as a fundamental tool for directed
change in education.

1. PRODUCT TESTING: the testing of products to describe
   their characteristics.

2. INSTITUTIONAL ACCOUNTING: the accumulation of data
   about an institution's operation—income, expenditures,
   costs per credit hour, faculty-student ratios, etc.

3. ACHIEVEMENT TESTING: The measurement of pupils'
   knowledge at the beginning and end of a course.

4. DIAGNOSIS AND ASSIGNMENT: the diagnosis of pupils'
   present knowledge and skills and the assignment of
   pupils to individualized instructional treatments.

5. SENSITIVITY TO GROUP PROCESS OR OPENESS TO CHANGE AND
   ADAPTATION: the procedures one used to facilitate
   change or innovation and the willingness to modify
   plans as they are carried out.

6. STUDY OF PERSONALITY DEVELOPMENT: the ways in which
   pupils' interests, attitudes, values, etc. change
   over time.

7. STUDY OF INSTRUCTION: the particular interactions
   between teachers and pupils, and the discovery that
   certain approaches work with some students and other
   approaches work with other students.

8. ACCREDITATION: the collection of data and its review
   by an accrediting agency.

9. DECISION-MAKING: the collection and use of information
   by administrators for decision-making.

Each of the evaluation components produces relevant and necessary information toward the quality control of instruction.

Decisions inherent in the evaluation process

The design of program evaluation is dependent upon several decisions which the evaluator must make. First of all, evaluation procedures are dependent upon the kinds of inforamtion needed by the program decision-makers. Some information is decision-oriented. This decision-oriented information is parochial in nature, refers back to the specificity of a particular program, does not demand generalizability, is often needed quickly and must be gathered in a short-period of time from a quick survey or tally. The decision-oriented information often refers to operations analysis or resource allocation. A second kind of information demanded in evaluation is conclusion-oriented. The conclusion-oriented information demands generalizability, relates the program to the external world, is gathered under research-type controls, and often concerns program processes and products. The conclusion-oriented information requires rigor in data gathering and analysis, should be planned in the design of the program, and is often the basis for the judgment of the worthiness of the program.

13

A second decision inherent in the evaluation process is related to the nature of the variables being measured. Some variables are more difficult to measure than are others. If the variables are internal in the subjects, are in the affective domain, require a degree of transfer to be seen, must be inferred, or are instable; several dimensions of measurement should be evoked. Without a multi-dimensional measurement scheme with these difficult variables, the credibility of the evaluation results is subject to question. As the nature of the variables approach the continuum of variables that can be more precisely defined and measured, the need for multi-dimensional measurement decreases. When the variables refer to fundamental skills, univariate measurement becomes more reasonable and there is less concern with the credibility of the results.

A third decision that the evaluator must make is in terms of size, scope and intensity of the evaluation. As the number of subjects increase, the manageability of the enterprise increase, but the model for analysis becomes more effective. With large samples, the probability basis for decision-making is improved but the flexibility of measurement is decreased. As a sample size decreases, there exists threats of lack of representativeness, decreases in reliability but also increased options

14

in measurement. As the scope of the evaluation is enlarged, the intensity and the specificity of the evaluation is decreased by the logistics problem alone.

A fourth decision which is faced by the evaluator relates to the audience. Differences in personnel seek different kinds of information from an evaluation. Information besides that sought by the reader must be stated in such a manner that confusion and misinterpretation do not occur, since the nature of evaluation data often established basis for improper interpretation. In the design of an evaluation the evaluator must be sensitive to each audience that will utilize the evaluation and develop the data needed by each audience source.

Concluding Remarks

An evaluation process in the educational setting introduces conditions for the adoption of scientific methodology. Evaluation in education should be an on-going process. The purpose of an evaluation system is to feed back information on all relevant aspects of the educational process on a continuous basis. This information feedback is a basis for directed change in education. Evaluation and the information feedback system provide a climate that facilitates quality control in education.

CHAPTER TWO

THE MASSACHUSETTS INFORMATION FEEDBACK SYSTEM:

A Four Phase Approach

William G. Conroy, Jr.

Massachusetts Department of Education

Rationale

In the design of the Massachusetts Information Feedback
System, consideration was made of both the positive and negative
aspects implied by state-wide evaluation. The design emphasizes
development of a comparable state-wide data pool, use of the
Research Coordinating Unit as a technical support component to
process and analyze the data, scheduled dissemination of feedback,
and programmatic research into instructional efforts to meet
state-wide and local needs. These activities must be carried on
while the system maintains local autonomy in administrative
decision-making, in curriculum design, and in school organization.

In order to be effective, state-wide evaluation must be a
partnership endeavor, allowing the state to fulfill the tasks
which it can do most effectively, namely curriculum design and
program modification. Evaluation results can be best used when
the information basis is broad, compatible, and impartial, and

17

when the results are anticipated, respected, and utilized in instructional decisions. Hence, data controlled by an agency external to the local school and based upon state-wide referents should provide a dependable basis for decisions. Modification resulting from this data, however, will only occur if and when it is desired, and its need realized in the local school.

Several unusual logistics problems are generated in the collection and by the control of state-wide information within a system which seeks to preserve local autonomy in curriculum design. The enormity of the management and development tasks dictate gradual implementation which include feasibility studies of the processes, evaluation of the total system from a smaller experimental system, training of the personnel in the development of objectives, the development of an adequate test battery, the design procedures for data collection, analysis and feedback, dissemination, and the training of personnel to produce the information and to utilize and interpret the results.

The Massachusetts Information Feedback System must then be thought of as an evaluation which seeks program improvement and modification rather than program condemnation. It must serve as an evaluation instrument, providing information on a state-wide basis and yet preserving local autonomy in curriculum design. The system is designed as a partnership endeavor between state agencies and local schools, and is conceived as a developing process.

To date, a four-phase system has been proposed and is being developed. Modification of this original design, however, is constantly under consideration as the system is implemented. The four phases have been designed to include the information essential to program management and instructional modification: namely program evaluation (more specifically product assessment), process-product evaluation, cost-effectiveness evaluation, and over-all evaluation (or impact study). Perhaps, the best method through which the design can be described is through the presentation of brief descriptions of each projected phase.

Phase I

Program Evaluation

This phase of the evaluation system will provide annual pre-test, post-test, and regressed gain scores for major objectives in every vocational or technical program in Massachusetts.* In each case, objectives will be determined by the local school for each vocational or technical program and tests designed to measure these objectives will be administered by the local schools, and analyzed by the Research Coordinating Unit on a pre-test, post-test basis. Program evaluation will provide both local and state

---

* See pages 46-47 for description of a regressed gain score.

means, standard deviations and estimated gain scores for grades 9-14. Information on attrition rates will also be included.

Phase II

Process-Product Evaluation

This phase of the evaluation would be conducted through a state-wide survey and through experimental study of a representative sample. Analyses will be performed to determine relationships between specific program variables and student success or non-success in career growth. This phase of the evaluation would consider students who are employed in fields for which they have been trained, students who are employed in fields unrelated to their training, and students who are unemployed. Examples of process variables might be: intensity of training experience, school facilities, teacher characteristics, school characteristics, etc. This phase of the evaluation would also examine relationships between degrees of success or non-success on program objectives and success or non-success in career growth. It would also examine relationships among schools offering similar objectives with different instructional processes in terms of regressed gain scores.

Phase III

Cost-Effectiveness Evaluation

This phase of the evaluation system also would be conducted

on a sampling basis and investigate the relationship between cost

and effectiveness of vocational-technical education. This phase

of the system would quantify and compare cost and effectiveness

of various vocational-technical educational programs. Cost-

effectiveness evaluation is a helpful planning tool for allocating

resources within vocational-technical education.

Phase IV

Over-all Evaluation of Vocational-Technical Education

This phase of the evaluation system would consider over-all

objectives of vocational-technical education and evidences of

total program impact on the state level. Data concerning over-all

programming, enrollment, and economic impact would be gathered

through state-wide surveys. The over-all objectives would be

assessed through study of random, representative samples. For

example, one over-all objective of vocational-technical education

might be: a large percentage of students who complete vocational-

technical educational programs should obtain and maintain

employment in fields for which they were trained and progress

satisfactorily in their career; a smaller percentage should

obtain and maintain employment in fields closely related to the

area in which they were trained and progress satisfactorily in

21

their fields. When compared with similar students who did not complete vocational-technical programs, students who completed vocational-technical programs should enjoy more job success. This difference should be significant. Other variables that might be considered could be: job satisfaction, citizenship behavior, social behavior, and general knowledge.

CHAPTER THREE

PHASE I: PROGRAM EVALUATION

Ray A. Johnson

Jim C. Fortune

University of Massachusetts

Rationale

The program evaluation has been designed to provide feedback
on the effectiveness of specific programs in achieving locally
entertained objectives. Two essential components of the product
assessment, the development of a file of behavorial objectives
for each program and the development of a test file for each
objective, constitute the major aspects of the program evaluation.
Each of these components will be described and procedures for
operation will be reviewed in the following sections of the guide.

A. WRITING BEHAVORIAL OBJECTIVES*

The purpose of this section is to bring about a common
understanding of the development of behavorial objectives. With
this goal in mind, it is only fitting that the author communicate
his objectives, so that the reader is able to measure the change
in behavior resulting from reading this section.

---

* This section of the chapter is a straight-forward, pedagogical
exposition of behavorial objectives and is the same material used
with instructors in Massachusetts. It was determined that inclu-
sion of this section, in its entirety, would precisely define
behavorial objective as we perceive them within the Massachusetts
Information Feedback System. (Ed.)

24

1. Given a list of one to twenty objectives, one identifies, with 100% accuracy, those objectives which are behavioral and those which are not.

2. When presented with a person whose area of vocational speciality is different from that of the reader, one develops behavioral objectives representing that person's field by questioning him about the pertinent information in his area of specialization.

Now that two examples of behavioral objectives have been presented, do they fit the following definition? BEHAVIORAL OBJECTIVES TELL WHAT IT IS THAT ONE WILL HAVE TO DO WHEN HE IS EVALUATED, THE CONDITIONS UNDER WHICH HE WILL HAVE TO PERFORM, AND THE LEVEL OR QUALITY OF PERFORMANCE EXPECTED.

The objectives of a course should contain as many statements, items, or examples as are necessary to describe the desired behavior of the student when he completes the course. Objectives of a course may be written in any form necessary which clearly states the instructional intent of the course.

This might represent an objective, "The student must understand the operation of a milling machine", but the

instructions still would have to go on to explain what is meant by "understand" by describing what the student will be expected to do to demonstrate the definition of "understand". Therefore, the objective, as stated, would not provide the necessary information as to what the student would be doing to demonstrate his achievements of the instructional intent.

The purpose of a behavioral objective in the context of this article is to make clear to teachers, students, and other interested persons what a person is expected to be able to do or perform when he finishes a course or program.

A well-written behavioral objective should say three things: It should say what it is that a student who has mastered the objective will do. It should say under what conditions the student will do this. It should say to what extent the student will do this. To put the matter in a single sentence, a well-written behavioral objective should specify under what conditions and to what extent a certain kind of student performance can be expected to take place.

Performance-conditions-extent. Let us consider, first, the word performance. Performance means doing. A student who performs something does something.

Here are two statements. Which one is expressed in terms

of student performance?

A. The student has a good understanding of electrical current.

B. The student measures electrical current.

Statement <u>B</u> tells what it is that the student will <u>do</u>. He will measure electrical current.

Statement <u>A</u> tells us that the student will have a good understanding of electrical current. But this is not very clear.

We cannot tell what the student is supposed to be able to do as a result of this understanding. The difficulty with using such verbs as understanding, appreciate, know, is that the performance is not directly observable. We have said that statement <u>B</u> is expressed in terms of student performance. Does this statement also set forth the conditions under which the performance is to take place?

No, it does not. For one thing, we cannot tell from our statement whether the student is to measure the current of an automobile, a house, a hi-fi amplifier, or a flashlight. Also, we do not know how he is to measure it; through mathematical computations, through the use of a meter, etc. Obviously, each set of conditions that one might develop is

27

substantially different from the rest, and will make its
own special demands upon the student who attempts to accom-
plish the objective.

Let us examine tr) more statements. Which one sets
forth the conditions under which a certain kind of performance
is to take place?

A. Given a circuit to operate a light bulb, the
   student indicates how to measure current with an
   ammeter.

B. The student will measure current at least 90%
   of the time when presented with electrical devices.

Statement A tells us that a circuit containing a light
bulb, and an ammeter will be used to set the conditions for
the demonstration of the student mastery.

Statement B offering us only the dubious clue of
"electrical devices", does not tell us enough. Our conditions
need to be defined more precisely than this. Therefore,
statement A would be the correct objective.

We come now to the matter of the extent and level of
performance. A well-written instructional objective establishes
acceptable minimum standards of achievement.

Look at Objective A: Given a detailed drawing containing
five cutting operations, the student performs them on a wood
lathe.

28

This objective is behavioral in that it states the desired performance that the student must exhibit and under what conditions, but it does not include an evaluative measure that would tell both the student and the teacher what is acceptable performance.

The objective should communicate how well the student performs when he has mastered the objective. Possibly, the most obvious way to indicate to what extent a student performs is to set a time limit. This criterion is easy to measure, but it does not always apply.

"The student disassembles a two-barrel carburetor in fifteen minutes." The statement "in fifteen minutes" leaves no doubt as to the extent of time the student has to perform the task. In some cases the time extent might be of little concern. The fact that the student can "disassemble a two-barrel carburetor" is fact enough to the accomplishment of a desired performance.

Whenever a minimal level of acceptance is desired, there should be an indicator that communicates to what extent the student should perform. The following are some examples that might provide some various ways of stating a criterion of acceptable performance.

Minimal Number

"When shown ten slides representing the Silver Culture of the N.E. Region of the U.S., the student matches the proper method of water and lumber management with at least nine slides."

Percentage

"Given the grass seeds that are indigenous to the N.E. Region of the U.S., the student categorizes them into the varieties of bent grass, blue grass, fescue, and rye with 90% accuracy."

Tolerance of Deviation

"On an engine lathe the student turns NC Thread #24 with the American Standard Association Class #2 fit."

In some instances his ability to perform a job with a certain degree of accuracy is pre-determined by the minimal acceptable requirements within a job area. In electricity, for example, the requirements are regulated by the National Electrical Code.

A minimal acceptable performance could relate to either an individual, an institution or a prerequisite requirement for job entry. Referring back to Objective A, it might read: Given a detailed drawing containing five cutting operations on an engine lathe, the student performs them within one-sixteenth of the specification set forth.

What should be included in a clear definition of an
educational objective? Sometimes there is confusion about
clarity and specificity. An instructor could end up with
several hundred objectives for one division within a
course which is probably too specific. Alternately, the
existence of only a few, general objectives lacks clarity
and fails to communicate enough specifics about the course
to enable measurement. Let us not be confused with specifying
the various teaching procedures or learning activities
as instructional objectives; they are not objectives. The
objectives should be stated at the level of behavior the
instructor intends for the student to acquire. For example,
a generalized behavior might be the ability to de-horn a
calf. The objective then could be stated as: Given a
caustic compound, the student de-horns a seven to fourteen
day old calf without burning the calf's skin or having
the horns grow back.

The emphasis is on the behavior to be learned by the
student rather than how the student acquired the ability to
perform the skills. As one can readily see, there are many
skills to be learned and a variety of knowledge that one
must accumulate to meet the objectives. However, the
objective does not specify or limit how one is to go about

31

acquiring the knowledge or skill. The issue is only how

the terminal behavior will demonstrate proficiency in

internalizing the skills and knowledge. Specifying the

kind of behavior to be developed and the kinds of content

involved is necessary to guide the selection of learning

experiences and to evaluate results. However, this

specification should avoid a degree of documentation which

leaves no alternatives for adaptive instruction.

A statement of a behavioral objective enables eval-

uation to be based on a student's performance and not on

an unknown element of appreciation, comprehension, or

understanding. This behavior which the student exhibits may

be anything from a manipulative demonstration skill to

writing an explanation of that skill.

After one knows what the necessary ingredients are in

a behavioral objective, he needs a systematic way to go

about arriving at the objectives. [During this pilot year

and because of the experimental nature of the information

feedback system, the Research Coordinating Unit is asking

for a representative sample of objectives.] An instructor

may take the already existing "K" blanks* and/or courses of

study and state the basic divisions which are the most

meaningful and cumulative of his instruction.

*"K" blank is a Massachusetts Task Analysis Form.

32

To increase communication, the following definitions have been accepted.

Course--the Vocation in which a student is enrolled (Automechanic-Horticulture)

Division--a field within a Vocation (Fuel system- Turf Management)

Unit--a task or breakdown within a division (Carburetion-Seeding)

This example of a conceptual matrix is suggested as an aid to help lay out a specific vocational field so that one might develop objectives at the proper level.

Grade Level or Level of Difficulty	Divisions					
	I	II	III	IV	V	etc.
9	5	5	5	5	5	
10	5	5	5	5	5	
11	5	5	5	5	5	
12	5	5	5	5	5	
Total	20	20	20	20	20	= 100 objectives

(develop objectives at each level in each division)

When a division of a course gives a person saleable skills, that division should be broken into behavioral objectives at the Unit Level. For example:

Turf Management is a division of Horticulture, but it is complete within itself. Therefore, the objectives would be developed at the Unit Levels of seeding-equipment, etc. The

33

Matrix would look like this.

Unit Levels

Grade Level	Seeding	Equipment	Budget	etc.
9	x	x	x	
10	x	x	x	
11	1,2,3,4,5	1,2,3,4	1,2	
12	x	x	x	

The objectives will be compiled and recorded on a form so that necessary and pertinent information about the objective can also be stated. For instance, an instructor desires a particular kind of testing procedure to be used, or knows of an instrument already in existence, it would be helpful to state that fact. Here is an example of the objective worksheet which is to be used this year.

OBJECTIVE WORKSHEET

School	Marshall High	Course	Electronics
Grade or Level	10th grade first level	Division	Circuitry
Instructor	Roger Tunks	Unit	Resistance

The following objective is stated in terms of observable student performance:

(Performance-Conditions-Extent)

1. Given five carbon resistors, a student will read the colorband and compute the resistance of four

out of five resistors with 100% accuracy.

Suggestion

    *Type of Test Desired: <u>Randomly select five 1/4 watt carbon</u>

_____

    *Type of Test Available:_____

_____

    Other Pertinent Information:_____

B. TESTING

Introduction

Since the evaluation and information feedback system
depends upon content designation by each individual school,
it is conceivable that each school would desire and develop
different behavioral objectives within each curriculum
division. Although this possibility exists, totally
different objectives for each curriculum division do not
seem probable; however, a variety of objectives within each
division must be anticipated in the design of the testing
process. The anticipation of unique objectives and the
variety of conditions existing in the sample schools demand

flexibility and continuous development of the testing process.
These conditions include different lengths of both classes
and school days, different administrative structures, varied
course sequences and organization, varied sources of related
courses and different resources, such as staffing patterns,
student composition and equipment. In order to establish
flexibility several formats of assessment will be utilized.
To assure continuous improvement of instrumentation a test
development staff will be located in the research coordinating
unit to identify and collect available instruments. These
include licensing tests, standardized tests, and certification
examinations currently being utilized in the field. This
will aid in developing tests specifically designed to assess
unique program objectives. This section will deal with
both the anticipated development of the testing process
and the planned analysis of the test data.

Format of Testing

Basically, the information feedback system will rely
on a pre-test, post-test format of testing, except when
pre-testing may be deemed inappropriate due to time
consumption or exceptional risks. The pre-test, post-test
format consists of criterion testing prior to the beginning
of the instruction in a curriculum unit and post-testing

with the same instrument after the instructional unit has been presented.

Since instructional units vary in duration and differ in the time of the school year in which they begin, the development of a manageable testing schedule is paramount in importance. The chronological schedule of testing will utilize the enrollment point immediately prior to the beginning of instruction and then will be post-tested at the end of that program level, such that the pre-test, post-test cycle will include the total program level or a yearly increment of instruction depending upon which is the shortest.

It is anticipated that the objectives of several units will be included in each program level and will be tested at the same or approximately the same time. Hence, during the first one or two days that a student attends classes for the program level in which he is enrolled, he will be pre-tested on the behavioral objectives for that program level. The same amount of time will be required for post-testing at the end of the program level. A maximum of two days of testing will be utilized for any program level.

Due to the nature of the objectives being developed and due to the nature of the curriculum, the pre-testing, post-testing

format on stated criteria will probably be inappropriate
in many cases. These cases will generally evolve when
the behavioral objectives demand time-consuming criterion
tests involving student performances or when neophytes
must be subjected to expensive or dangerous machinery in
order to be pre-tested. Both of these situations tend to
require performances which can be simulated for pre-testing.
Hence, a simulation test designed to provide entry diagnosis
of fundamental skills or knowledge will be used instead
of the criterion-centered pre-test.

Two variations of testing are conceivable within the
simulated pre-test,criterion-centered post-test format. In
the first variation, the simulation test will contain
enough fidelity to the post-test, that the post-test per-
formance will demonstrate criterion attainment from a
baseline defined by the simulation pre-test. Hence, criterion
attainment becomes the evidence of gain and degree of such
attainment constitutes the instructional product.

In the second variation, the simulation pre-test would
not be interpreted as representative of adequate baseline
information, since an assumption could be made that criterion
performance may be influenced but not depend totally on

successful performance on the simulation tests. Two types
of post-testing become essential; both the criterion-related
post-test and the simulation test would be given at the end
of the program level. Corrected gain scores on the simulation
test would be used to ascertain change and the criterion-
related post-test would demonstrate the accomplishment of
the objective. Using the criterion-related test, progress
would be shown through the post-test format only, since
entry into the course assumes a zero beginning point. (In
this case, success on simulated pre-tests would indicate the
need to either allow the student to be pre-tested on the
criterion-related test or to pursue a more advanced objective.)
In retesting with a simulated test, changes in prerequisite
behavior of the student could be shown.

In both formats the basic objective is to estimate
the results of instruction designed to develop a specific
objective. Although the formats tend to produce different
kinds of evidence of instructional results, the evidence
produced can be used in the same manner in the decision-
making required to improve and modify instructional tech-
niques.

Test Administration

The three facilitators at each sample school will be asked
to coordinate the testing formats.* The specifics of test
administration will be expected to vary among the schools; how-
ever, the testing formats will be held constant. Testing
schedules for each program level will be developed for each
school and coordinated between schools so that comparable in-
structional units of time are maintained. Both the pre-tests
and post-tests will be administered according to the administra-
tion instructions included with each test at each school and the
completed tests will be sent to the Research Coordinating Unit
for analysis.

Test Development

One of the major activities of the Research Coordinating
Unit is the development of a test battery for measurement of
the behavioral objectives. This test battery will be composed
of a collection of existing tests and licensing examinations
which appropriately meet the testing requirements of the sys-
tem, the development of tests in areas where no suitable tests
can be found, and the specification and standardization of
rating procedures for performance tests. After a coding system
has been utilized to categorize the objectives, a similar cross

---

* Applies to pilot study only.

indexed file of measuring instruments will be set up. This file will contain the testing instruments and a historical profile on each test.

A brief discussion of the information which the historical profile will contain is in order. The profile will provide four kinds of information about the test, namely; (1) information on proper usage of the test including the objectives for which the test is appropriate, when and with what testing format the test should be used, and a record of test usage; (2) information upon the effectiveness of the test including reliability and validity evidence; (3) information upon how the test should be administered so as to assure comparability of testing conditions; and (4) information upon the measurement procedure used by the test so that proper methods of data analysis can be utilized.

The information on proper usage of the test will include a list of objectives for which the test has been designed, the curriculum division and level for which the objectives were written, the testing format within which the test should be used, the source of the test, and a record of who, when, and where the test has been used.

The information upon the effectiveness of the test will include several kinds of evidence upon reliability and validity, some of which is essential for the computation of estimated gain scores and some of which is needed for proper interpretation of the results. Reliability refers to the consistency of test performance during repeated measures. A test can be said to be realiable when very similar results are obtained for repeated applications of the test. Three types of reliability will be collected for each test. Test, retest reliability (the consistency of the test results for two administrations of the same test to the same group of students) is essential for the computation of estimated gain scores and will be studied for each instrument. Split-half reliability (the degree of consistency of measurement between the odd-items and the even-items of a test) and parallel-form reliability (the degree of consistency of measurement between comparable forms of a test) will be reported also.

Validity is the criterion for adequacy of a test which deals with the degree to which a test measures what it has been intended to measure. Evidence of three forms of validity will eventually be developed on each test. The first form of validity that

is relevant is content validity which requires judgment concerning how well the test samples the behavior which it is supposed to test. Evidence of concurrent validity (how well the test corresponds or correlates with a second criterion measure taken at the same time) and evidence of predictive validity (how well the test corresponds to or correlates with a future criterion measure) will also be developed and included in the profile.

Information upon how the test should be administered will include the time required for proper administration of the test, the supplies and equipment needed to administer the test, the setting required for testing, the directions which must be given to properly administer the test, and the forms upon which test data is to be reported.

The information upon the measurement procedure used by the test is necessary to determine proper statistical treatment of the data. Four forms of measurement are used in testing and each form places demands upon the type of analysis that can be made on the data. In nominal measurement, labels are assigned to objects so that these objects can be placed into categories. No attempt is made to place value or size to the categories. In ordinal measurement,

objects are ranked in regard to a characteristic, but there
is also an attempt to estimate the difference in the char-
acteristic between ranks through the establishment of
intervals which approximate the magnitude of existence of the
characteristic. In ratio measurement, both the ranking
and the assignment of magnitude by equal intervals which
occur in interval measurement are utilized, but an additional
refinement of measurement in terms of an "absolute zero" is
added where there is no evidence of the characteristic
measured.

Analysis and Reporting of Test Results

Generally, the evaluation feedback system is concerned
with a descriptive comparison of specific class distributions
to the parallel statewide distributions on a common objective.
The feedback report form will provide the statistics for
these basic comparisons, but the comparison and analysis
of its meaning must be made by the local educator who has
the most complete knowledge through which decisions can be
made. The report form will include the means, the standard
deviations, the number of scores analyzed, and the estimated
gain scores on an objective for the class and for every other
class entertaining that objective.

For the clarity of interpretation of the data, the meaning of the statistics reported back should be explored. First of all, it should be noted that the number attempting the objective and the number satisfactorily completing the objective is in itself meaningful and usable data. Differences from the number entered and the number finished for a program and for the entire state provides a basis through which attrition rates can be compared.

Also, reported in the form is the mean and standard deviation on both the class and the state. These two statistics serve to describe the distributions of the two sets of scores. The mean (mean equals sum of scores divided by the number of scores) used is the arithmetic average of the scores made by the group. The means serve as measures of central tendency and describe the performance level of the class and of the state as a whole. The standard deviation indicates how the class scores were spread around the mean. Hence, by use of the mean and standard deviation, the report presents a visual picture on what the average performance was and how members of the class performed in regard to that mean. By showing both average performance and distribution

of scores around this performance for both state and local
programs on a pre-test post-test basis, change of the local
group can be compared to change of the state group. The
standard deviation can be thought of as the average distance
of the individual class scores from the mean, although a
squaring process and then the square root of the summated
results is used to get rid of the sign problems created by
the existence of subjects falling on both sides of the
mean. Hence, the standard deviation represents the average
distance of the scores from the means, but it is calculated
by taking the square root of the mean-squared deviation
from the mean.

The gain scores that will be reported have been called
"adjusted" or "estimated" gain scores since a statistical
correction has been made to account for chance variation
and to remove the difficulty of unequal rates of change for
extremely high and extremely low scores. The correction
will be achieved by using the Lord model (See in C.W. Harris-
Problems in Measuring Change, University of Wisconsin Press,
1967.) predicting gain by regression of the pre-tests on
the post-test. The Lord formula states that $\hat{G}$ (which is
individual student gain) equals G (which is the difference

of pre-test and post-test means) plus b_{Gxy} (where b_{Gxy} is

the partial regression coefficient defined by the formula

$$b_{Gxy} = \frac{(1-r_{yy'})r_{xy}s_y/s_x - r_{xx'} + r^2_{xy}}{1-r^2_{xy}}$$

where $r_{yy'}$ is test, re-test reliability on the post-test

measure, r_{xy} is the correlation between pre-test and post-

test scores, s_y and s_x are standard deviations of x and y,

and $r_{xx'}$ is the test, re-test reliability for the pre-

test) times the deviation of x from the mean $\overline{x}$ minus b_{Gxy}

(where $b_{Gxy} = \dfrac{r_{yy'} - r^2_{xy} - (1-r_{xx'})s_x r_{xy}/s_y}{1-r^2_{xy}}$

from the mean of the post-test. Hence, an estimated gain

for each student can be computed and an average of these

is the reported gain score. This estimated gain score

would then lead to comparisons between the class and the

state on a group level. Upon request, however, gain scores

for individual students can be obtained for more specific

analysis of instruction.

CHAPTER FOUR

PHASE II: PROCESS-PRODUCT EVALUATION

David C. Berliner

University of Massachusetts

Overview

When the process dimension in evaluation is discussed,
an attempt is made to specify some distinct and defineable
factors that may be separated from the totality of events
which impinge upon a student as he goes through a formalized
learning program. Only those factors which may have
influenced the final product, a skilled student, are to
be recorded. One should try to record, in some way, those
significant events in the learning program which may have
contributed to the quality of the product. A conceptual-
ization can be illustrated in Figure 1.

Figure 1. Flow Diagram of Educational System

FORMAL LEARNING PROGRAM

Entry   S   O   I   A   Exit

Naive Student →   STRUCTURAL   ORGANIZATIONAL   INSTRUCTIONAL   ATTITUDINAL   Performance Criteria → Skilled Manpower

Raw Material   Process   Product
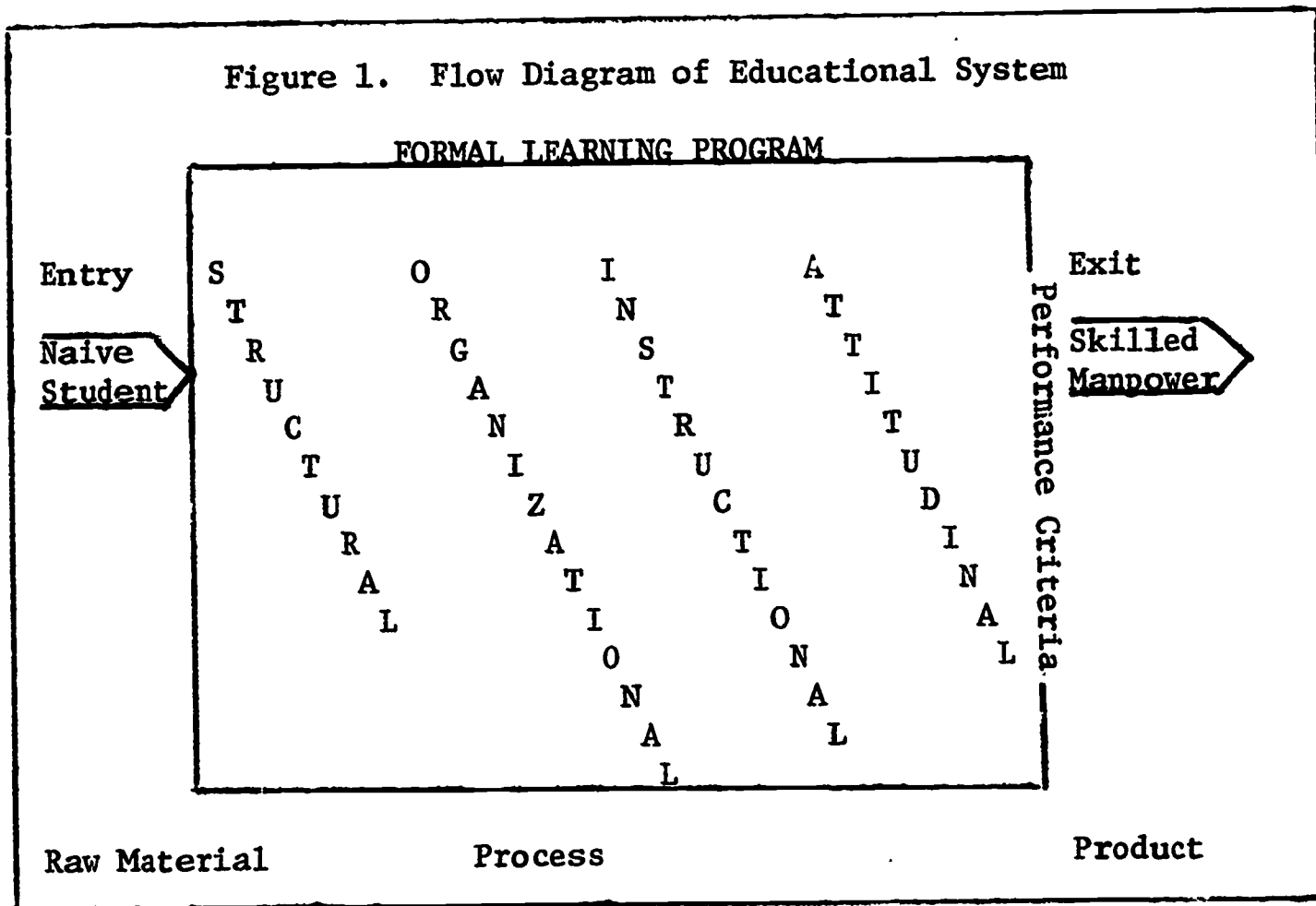
It is possible to conceive of untrained or naive students entering some formal learning program and leaving after 1 day, 6 weeks, or 4 years, as a skilled person in his field, be it music, English, or auto mechanics. While in the program, certain structural considerations might be thought of as having an influence on the quality of the final product which is turned out. That is, the quality of the learner's experience may be effected by such things as the size of the building, the number of

units of a certain piece of equipment, etc. These types
of variables may be categorized under the heading of struc-
tural considerations when developing ideas about important
process components. Further discussion of these ideas
will appear below.

Just as certain structural considerations define
certain opportunities and limitations in learning, so do
the organizational developments which occur. Organizational
variables which might impinge on the student as he goes
through a learning program are size of class and type of
degree program (vocational, general, academic). All matters
that are related to the way the learning process is organized
within certain kinds of structures are included in this
category and elaborated on below.

Figure 1 notes also the instructional considerations
occurring within certain structural and organizational
patterns. Included for consideration in this category
are such factors as ability rating of the instructor,
preparation of instructional material and degree of individ-
ualization of instruction. In short, this category will
be used to conceptualize the variables in interpersonal

51

teacher-student contact which may bear on the quality of
the product.

Finally, one may conceive of the development of certain
attitudes which students develop as a direct result of,
or because of interaction with the structure, organization,
and instruction to which they are exposed. The attitude
toward their school, and towards their schooling in general,
would be important information to have. Such things as
the student's feelings of competency, and feelings about
his job prospects may be considered. Virtually, any
attitudes toward school, self, and society might come
under study.

The preceding discussion of structure, organization,
instruction, and attitudes all may influence performance
at stated criterion levels. The evaluation of performance
through a statement of behavioral objectives has been dis-
cussed earlier in this document. The variables mentioned
in this section are variables which may influence that
performance level of students and are, therefore, important
because of their effect on our product...skilled manpower.
A more detailed discussion of each category and the variables

to be considered, therein, is provided below.

Structural Considerations

The intent of defining a category called structural
considerations is to pick out those aspects of the forma-
lized learning process which are either invarient, or at
least hard enough to change so that they appear invarient.
Thus, we are distinguishing between the more static aspects
of an educational system and the more dynamic aspects.
In the latter case, we would be discussing organizational
considerations. The organizational aspects of a system
are more fluid and subject to change than are the structural
aspects, which more nearly represent the "givens" of the
situation, thus making this the hardest-to-change aspect
of the system.

In considering structural aspects which may effect
the product, one might want to note something about the
facilities and equipment available, i.e., the resources
brought to bear on the learning process. One might ask
about the buildings; their size; student capacity; age;
and dollar value of the plant. Under this category,
information about the available equipment might include:

number of pieces of equipment of a specialized type, i.e.,
V-8 engines, lathes, and oscilloscopes; dollar value of
equipment; age of equipment; etc. One could include, within
this category, information on the socio-economic and demo-
graphic characteristics with which a school must concern
itself, for example: the neighborhood within which the
school lies; neighborhoods from which the school draws
students; whether the vocational training is occurring
within an industrial community, or whether vocational educ-
ation is occurring within a rural or suburban community.
The indebtedness of the district might be questioned; tax
rate for schooling; and per pupil expenditures in vocational
and academic programs. These could then be used as indices
of a district's commitment to educational quality.

The category of structural considerations is intended
to aid the reader in conceptualizing those difficult to change,
or invarient aspects of an educational system. Disagreement
or agreement with the examples is not important. Certainly,
no attempt was made to be exhaustive. A set of variables
which may influence the product have been illustrated and
these variables seemed to group naturally under a heading
called structural considerations.

Organizational Considerations

Under the category of organizational considerations in process documentation, those aspects of learning processes, which are fluid and subject to much more change than the variables listed under the structural dimension are examined. For example, one might include among organizational processes the size of classes; the number of periods per day; whether the school is using open lab or closed lab concepts; etc. All the above are conscious decisions about the way a school should be organized and, therefore, subject to change. The totality of these organizational considerations would represent at any one time the belief system of the faculty concerning the best way to organize the learning process. Under the heading of organization, one then can include such things as: number of business advisory groups; size of those groups; time spent in meetings per month; and all other aspects of the advisory function which may have a direct bearing on the courses taught in a vocational program and the skills to be mastered in that program. One might want to include under organizational considerations such things as pupil-teacher ratio; the qualifications of

pupils in academic, general, and vocational programs; the

number of units or credits necessary for the completion

of a course of study; and the number of vocational electives

allowed.

The point in specifying an organizational category is to

be able to list those aspects of a formal learning process

which, through vote or tradition, have led to certain

educational practices which may effect the product in which

one is interested.

Instructional Considerations

Under this category variables will be documented which are

involved in the instructional process and which may in some

way effect the degree of proficiency which a student demon-

strates on completion of his training. Here, one might

wish to consider such things as the qualifications of the

instructor; highest degree held; years of schooling; years

of practical experience; positions held in industry; in-

service programs in education; and professional affiliations.

Perhaps, one would want also to record such things as salaries

paid to instructors; differences in vocational educator's pay

compared to the scales used in the academic areas; and certi-

fications held or required. Support for the instructional

process may be recorded under this category and include such things as number of computer terminals; library facilities; audio-visual equipment available; and number of screens. The point, of course, is to pick out those variables which may effect the quality of the product.

In addition to the support for instruction and background of the instructors, it might well be important to employ measures which tap the interpersonal dimension in teaching. Is the teacher warm, friendly, clear in explanation, well-prepared, and fair? The characteristics of an instructor or the "climate" created by an instructional staff can certainly be expected to impinge on the student as he goes through an educational system.

Attitudinal Considerations

Given a certain structure and organization for an educational system within which a certain kind of instruction takes place, student attitudes of varying degrees can be expected to develop on many divergent topics. To tap such information would aid in the overall program design to obtain data which leads to modification as well as evaluation. Thus, under this category one might consider important the student's feeling of personal worth; does the

57

self-image of the student reflect confidence and pride?
what attitudes exist toward certain courses? toward
instructors? toward school in general? toward their school
in particular? Information about reality testing by the
student such as his feelings about the pay he is likely
to receive; attitudes toward unions; toward job stability
and mobility; toward his own prospects to obtain employment
might be documented.

- Since people hold opinions, attitudes and beliefs
about myriad events, some decisions about which of these
are important must be made. These decisions, as with all
other decisions about which variables to examine, must be
made on the basis of known or hypothesized effects on the
quality of the final product of the system. Described
below are methods for obtaining information about process
variables deemed important in evaluating an education system.

Procedures

The goal of process-product evaluation is the detection
of relationships between what the student experiences and
what his skills are. Thus far product has been discussed as
if it were uniform, but in fact there exists excellent, fair
and poor products in an educational system. There are

products who become regularly employed, irregularly employed, and unemployed. There are products who are employed in capacities that use the skills in which they were trained and others who are employed in areas for which they were untrained. These particular type of outcome variables, the product dimensions, may be related to certain antecedent conditions such as the structure, organization, instruction, or attitudes involved in the educational process.

There are many ways to record process information which may be useful in determining some relationships with outcome measures. One of these is the process log.

A process log is a continuing record of important events which may be related to performance. It is recorded by an instructor shortly after some instruction has taken place. The log might be broken up into records of the teaching that has occurred for each statement of performance criteria which had been set. Thus, if the performance criteria was an adjustment of a two barrelled carburetor to the required idling speed for a specific engine model, a teacher might record as an entry in the process log information about how the material was taught, as shown in Table 1.

Table 1.   Process Log Example

Date	Performance Criterion	Log Entry
3/24/70	#821 Carburetor Adjustment to idling speed.	Carburetor adjustment taught by lecture, no laboratory practice until one week after lecture. No model used for demonstration. Diagram with labelled parts was used during lecture. Lecture took 50 minutes.

This kind of entry--especially when evaluated against teaching techniques which either included a demonstration of an actual carburetor adjustment, or which placed unadjusted carburetors in front of each student to adjust as the instructor worked, or which used magnification shots over closed circuit television to show the procedures, etc.--then becomes the documentation of the process used and serves to ferret out for an investigator those procedures that work from those procedures that do not work, insofar as obtaining optimum performance for a stated objective.

Essentially, then, a process log is a documentation of methods used to achieve certain ends. In addition to the uses of the log noted in the illustration shown, a log can be used to record time in laboratory, number of students in the group, fixed equipment costs, amount of

equipment available, length of class time, etc. Thus, for
the start of a unit on the use of the oscilloscope, the
process log could include statements about structure
(number of oscilloscopes available, performance level of
students); statements about organization (class periods
of 50 minutes, 27 students all seniors); statements about
instruction (I hold a Navy electronics technicians certi-
ficate, no movies available, I feel good teaching this
material and am well-prepared); and statements about
attitudes (students seem eager to learn this, only a small
number are showing interest, few smiles are evident, etc.).
The log then would provide the instructor's view of the
learning process: what he did, what he used, what he
noticed, and with whom he worked are all data to be
recorded and preserved, for later distillation and examina-
tion.

Other ways to document process variables exist. For
some variables of interest other techniques would make for
better documentation. These techniques include rating
scales and check lists. The simple rating scale makes
the person doing the rating the measuring instrument. All

the scale does is systematize that measurement along a
continuum. For example, using the categories described above,
we may use the following kinds of rating scales:

Structural

This school building is:

unsatisfactory, a hinderance to learning	poor but service- able	about average in condition	good, should aid learning	excellent, completely satisfactory

Organizational

Class size for this learning sequence was:

too large, injurious to learning	too large	about right	smaller than need be	small enough for extensive individual attention

Instructional

Class participation:

almost never occurs	occurs with urging	occurs regularly	occurs quite frequently	always occurs

Attitudinal

Students appear to be:

always inter- ested and cooperative	usually interested and coop- erative	average in their inter- est and coop- eration	sometimes interested and coop- erative	hardly ever interested and coop- erative

62

These kinds of rating scales can be made up for variables of interest and used with individual instructional segments for meeting performance criteria, or for longer units, or for whole curricula. They can be used repetitively, singly, or in the case of some scales even by random schedules in order to monitor the systems under study. The documentation of attitudinal considerations should utilize the students as well as the instructor as the measuring instrument. The students provide valuable feedback from a different and important point of view.

The use of checklists can also play an important part in the documentation of process variables. This simple device allows a respondent one of two responses, agree or disagree (dislike or like) as evidenced by either a check mark or a lack of a check mark given to a particular statement. Some examples follow:

Check those items with which you agree:

_____1. Our facilities are adequate.

_____2. There is too much moving around during the school day.

_____3. Teacher-student relations are pretty good at this school.

_____4. Students are not well-prepared for employment when they leave here.

The above example includes a checklist item from
each of the categories previously described. Hundreds of
checklist items could be developed and the instructor
and/or the students could fill out such lists.

The documentation of process variables by way of
rating scales will allow an analyst to average the scale's
scores (provided equal appearing intervals were used)
and meaningful pre-post comparisons could then be made.
In the case of the checklist, information would be tabu-
lated in precent of respondents who agree with an item.
Changes over time in percentages may also be evaluated
statistically for significance if relevant program changes
have taken place between the two administrations of the
checklists.

Still another, and perhaps, the most popular form of
documenting process information is through the use of
standard questionaire items. Information about variables
in each of the categories noted might be obtained through
the use of a questionaire which provides all the background
information necessary to interpret, in a more meaningful
way, many of the entries in a process log. That is, to
know that an objective concerned about welding was taught

64

by film is useful. To know also that the student body at a school is primarily from families receiving welfare would make our information base about the learning associated with the film that much greater. Extensive questionaire development might accomplish that task. Some examples of questions used by the U.S. Office of Education to examine process are provided. Some item formats are like the rating scale and checklists.

Examples of questions most easily answered by district personnel might include:

A. Which of the following best describes the location of this school district?

   ( ) a. Large city (over 500,000)
   ( ) b. Suburb of a large city
   ( ) c. Rural area near a large city
   ( ) d. Middle-size city (50,000-500,000)
   ( ) e. Suburb of a middle-size city
   ( ) f. Rural area near middle-size city
   ( ) g. Small city or town (less than 50,000)
   ( ) h. Rural area, not near a large or middle-size city

B. Indicate the current starting annual salary in your school district for a beginning credentialed high school teacher with a bachelor's degree. $_____.

C. Which of the following types of personnel in your district have participated in in-service training programs since June, 1968? (Mark the percent of that group who participated.)

( ) a. Regular classroom teachers
( ) b. Special teachers (e.g. remedial reading teachers)
( ) c. Teacher aides
( ) d. Other professional personnel (e.g. health
       or guidance personnel)
( ) e. School principals

Other questions most easily answered by a principal at a

particular school include:

A. Grades in this school are ( ) K-6 ( ) K-12 ( ) 10-12
                             ( ) 1-6 ( ) 1-12 ( ) 7-8
                             ( ) K-8 ( ) 7-9   ( ) Other
                             ( ) 1-8 ( ) 9-12   (list)

                             _____

B. ADA for January was _____(total for all grades)

C. The number of new enrollees last year was _____.

D. The number of transfers or withdrawals last year was ____.

E. How old is the main classroom building of your
   school plant?

   ( ) a. Less than 1 year old   ( ) e. 20-29 years old
   ( ) b. 1-4 years old   ( ) f. 30-39 years old
   ( ) c. 5-9 years old   ( ) g. 40 years or older
   ( ) d. 10-19 years old

F. Indicate in FULL-TIME EQUIVALENTS the number of the
   following types of professional staff members
   available to pupils in your school:

   a. Regular Classroom Teachers (academic subjects)_____.
   b. Regular Classroom Teachers (vocational subjects)_____.
   c. Special Instructional Personnel (speech, physical
      education, art, music, reading, etc.)_____.
   d. Health Personnel (school nurse, school physician,
      etc.)_____.
   e. Psychological Personnel (social workers, counselors,
      school psychologist)_____.
   f. Media Specialists_____.
   g. Paid Community Personnel_____.
   h. Volunteer Community Personnel_____.

66

G. In your opinion, which of the following best
    describes the dwellings in the immediate area
    of this pupil's home?

    ( ) a. Well-kept single family houses
    ( ) b. Well-kept multi-family dwellings
    ( ) c. Run-down single family houses
    ( ) d. Run-down multi-family dwellings
    ( ) e. Don't know

H. Sometimes there are occasions when special problems
    such as natural disasters, epidemics, teacher strikes,
    etc. have long and prolonged effects on instruction.

    Have there been any significant acts, affecting
    your school, such as those listed above?

    ( ) Yes
    ( ) No

    If yes, how many weeks was instruction interrupted?

    ( ) One week
    ( ) Two weeks
    ( ) Three weeks
    ( ) Four or more weeks

I. Are students grouped in classes

    ( ) Heterogenously
    ( ) Homogenously

J. What percentage of the pupils in this school are
    members of families whose HEAD OF HOUSEHOLD did not
    complete the 8th grade?

    ( ) None ( ) 1-10% ( ) 11-25% ( ) 26-50%
    ( ) 51-75% ( ) 76-90% ( ) 91-100%

K. What percentage of the pupils in this school are
    members of families whose primary means of support
    is a public welfare program?

    ( ) None ( ) 1-10% ( ) 11-25% ( ) 26-50%
    ( ) 51-75% ( ) 76-90% ( ) 91-100%

Certain questions are perhaps best filled out by the

teacher.   Examples of such questions are:

L. What is your sex?
    ( ) Male   ( ) Female

M. How many years of full-time teaching experience
    (public and non-public) including this year, have
    you had?

    ( ) a. One year or less
    ( ) b. More than 1 year but less than 3 years
    ( ) c. At least 3 years but less than 6 years
    ( ) d. At least 6 years but less than 10 years
    ( ) e. 10 years or more

N. How many years, including this year, have you
    taught in this school?

    ( ) a. One year or less
    ( ) b. More than 1 year but less than 3 years
    ( ) c. At least 3 years but less than 6 years
    ( ) d. At least 6 years but less than 10 years
    ( ) e. 10 years or more

O. Do you reside within the attendance area or
    neighborhood of this school?

    ( ) Yes   ( ) No

P. Are you a member of one of these NATIONAL MINORITY
    GROUPS?

    ( ) a. Yes, American Indian   Yes, Spanish-surnamed
    ( ) b. Yes, Negro   American of:
    ( ) c. Yes, Oriental   ( ) d. Cuban Descent
                                  ( ) e. Mexican Descent
                                  ( ) f. Puerto Rican Descent
                                  ( ) g. Spanish Descent

    ( ) h. No

Q. Did you choose to teach at this school this year?

( ) a. Yes, from among many alternatives
( ) b. Yes, but the choices were few
( ) c. No, I was assigned here
( ) d. No, there is only one elementary school
in this district

R. Would you prefer to be teaching in a different
type of school?

( ) a. Prefer a different type of school
( ) b. Prefer another of this type
( ) c. Satisfied here
( ) d. Greatly prefer this school
( ) e. Don't know

S. Estimate the proportion of pupils in your class who
come from families in which the head of the household
is employed at the following level: (Mark one answer
in each line.)

	None	1-10%	11-25%	26-50%	51-75%	76-90%	91-100%
Children of professional or managerial workers	( )	( )	( )	( )	( )	( )	( )
Children of skilled workers	( )	( )	( )	( )	( )	( )	( )
Children of semi-skilled workers	( )	( )	( )	( )	( )	( )	( )
Children of non-skilled workers and laborers	( )	( )	( )	( )	( )	( )	( )
Children of agricultural workers	( )	( )	( )	( )	( )	( )	( )
Children of disadvantaged--welfare or unemployed	( )	( )	( )	( )	( )	( )	( )

T. What proportion of the pupils in your class are members of the following NATIONAL MINORITY GROUPS? (Mark one answer in each line.)

	None	1-10%	11-25%	26-50%	51-75%	76-90%	91-100%
American Indian	( )	( )	( )	( )	( )	( )	( )
Negro	( )	( )	( )	( )	( )	( )	( )
Oriental	( )	( )	( )	( )	( )	( )	( )
Spanish-surnamed American of:							
Cuban Descent	( )	( )	( )	( )	( )	( )	( )
Mexican Descent	( )	( )	( )	( )	( )	( )	( )
Puerto Rican Descent	( )	( )	( )	( )	( )	( )	( )
Spanish Descent	( )	( )	( )	( )	( )	( )	( )

U. According to your own judgment, what proportion of the pupils in your class are generally performing below grade level?

( ) a. None   (.) e. 51-75%
( ) b. 1-10%   ( ) f. 76-90%
( ) c. 11-25%   ( ) g. 91-100%
( ) d. 26-50%

V. What are the average number of minutes per INSTRUCTIONAL PERIOD, number of classes per week, and number of weeks per year spent in instruction in your class in your subject area?

My area is_____.

a. Number of minutes per instructional period: (check one)

1. 1-20   ( )   5. 61-75   ( )
2. 21-30   ( )   6. 76-90   ( )
3. 31-45   ( )   7. More than
4. 46-60   ( )   90   ( )

b. Number of classes per week:

1. less than 2   ( )
2. 2 but less than 4   ( )
3. 4 but less than 6   ( )
4. 6 but less than 8   ( )
5. 8 but less than 10   ( )
6. 10 or more   ( )

c. Number of weeks per year:

1. 0-6   ( )
2. 7-12   ( )
3. 13-24   ( )
4. 25-30   ( )
5. 31 or more   ( )

W. The list below is a list of educational ends or values. There is no right or wrong way to complete this; it is a description of values which a number of educators believe to be important ends of education. Please list your personal preference for those values by marking in Column 1 those values you think are most important, marking in column 2 those values you think are somewhat important, but not most important, and marking in Column 3 those you think are least important. (Mark one answer for each item.)

		1	2	3
a.	Development of obedience	( )	( )	( )
b.	Development of physical and motor skills	( )	( )	( )
c.	Development of respect for the rights of others	( )	( )	( )
d.	Development of good manners	( )	( )	( )
e.	Development of creativity and self-expression	( )	( )	( )
f.	Improvement of one's self-concept	( )	( )	( )
g.	Development of intellectual abilities	( )	( )	( )
h.	Development of personal responsibility for property and materials	( )	( )	( )
i.	Development of vocational interests	( )	( )	( )

Some items which may be appropriate for the pupils themselves to answer or which a teacher could answer about students are:

X. What is the average number of days absent per pupil per month?

( ) 0-1/2   ( ) 1/2-1   ( ) 2-3   ( ) 4-5   ( ) 6 or more

Are these absences primarily due to illness?   ( )Yes   ( )No

Y. If these pupils are given a difficult task to complete, on their own without prescribed time limits, they usually:

( ) a. Give up without trying
( ) b. Start the problem, but give up rapidly
( ) c. Persist for some time, but do not complete the task
( ) d. Complete the task that was assigned

Z. Items for the student directly:

Who lives with you at your home?

a. Father	( ) Yes	( ) No		
b. Mother	( ) Yes	( ) No		
c. Grandmother	( ) Yes, one	( ) Yes, both	( ) No	
d. Grandfather	( ) Yes, one	( ) Yes, both	( ) No	

e. Number of Brothers ( )0 ( )1 ( )2 ( )3 ( )4 ( )5 or more
f. Number of Sisters ( )0 ( )1 ( )2 ( )3 ( )4 ( )5 or more
g. Others ( )0 ( )1 ( )2 ( )3 ( )4 ( )5 or more

Which would make your parents happier to hear?

( ) a. You were doing very well in your school work (e.g. reading)
( ) b. You were behaving very well during school

Which would make your parents unhappier to hear?

( ) a. You were doing very badly in your school work (e.g. reading)
( ) b. You were acting very badly during school

Do you have a job after school?   ( ) Yes ( ) No

72

Summary

Though no attempt was made to break these items up
into the categories discussed above and no attempt to
exhaust relevant questions was intended, the examples
given should serve as a stimulus for a creation of
questions (rating scales and checklists) and for prepa-
ration to record important data in a process log. Al the
above material was meant to be illustrative of an approach
to documenting the important process variables. Flaws
might appear in the examples and the categories which
have been provided. Good! One should revise the formats
to his own liking and add other relevant important variables
to document materials on the process of education. It is
wise not to use everything presented by way of illustration
in a verbatim manner. Variables should be selected to
help educators achieve their goals.

What is important to realize is that improvement of
teaching techniques and educational systems will only be
possible when it is known which techniques are tried, with
whom, in what kind of setting, under what kind of conditions,
using what kind of equipment and with what sort of observable

CHAPTER FIVE

PHASE III: COST-EFFECTIVENESS EVALUATION

Roy Forbes

University of Massachusetts

75

Cost benefit and cost effectiveness analyses provide deci-
sion makers with financial data which can be used in planning
the efficient operation of manufacturing and service systems.
The managers of systems must decide how to invest resources to
maximize returns. They must consider both short and long range
goals. They must evaluate current plans and operations to deter-
mine if investments will lead toward the anticipated results.
They must also evaluate current operations to insure the effec-
tive use of resources. Cost analyses aid managers in these
responsibilities.

The professional educator is also expected to make policy
and operation decisions which are partially based on financial
considerations. Cost analysis is an available tool for their
use, but the complexities of educational systems does not permit
simple application of these tools.

There are many problems which limit the use of these

techniques. The lack of a measure of the quality of the output of the system is the largest. However, before the quality can be measured, the output must be defined. A basic question must be answered. "What is the product (the output) of an educational system?" Cost studies usually tend to ignore this problem.

Product (output) is normally measured as academic achievement or in terms of the economic contributions which an individual is expected to make to society. Are academic achievement and dollars suitable measures of the output of an educational system?

The main purpose of this section of product evaluation is to present a suggested approach to cost analysis which is based on the relationship of institutional objectives with instructional programs and related costs. A basic assumption of the proposed technique is that measurable institutional objectives will specify the desired output of the system.

First, however, is a review of the classical cost benefit analysis of an educational system.

The comparison of cost--the investment of resources--and financial benefits of a system is a cost-benefit analysis. Analyzing the financial benefits of the investment of resources provides data for planning efficient programs. Hence, cost-

benefit analyses of educational systems are highly desirable.

There are many interrelated variables which need consideration in a cost-benefit analysis of an educational system. Financial benefits may be divided into four dichotomic classifications: individual-society, measurable-non-measurable, direct-indirect, and positive-negative.

The first classification distinguishes between benefits accountable to the individual, e.g., increased earning power, and benefits of society, e.g., an increased personal income tax base. These two examples are also measurable benefits. A non-measurable benefit (or one difficult to measure) is the effectiveness of the individual as a contributing member to the political and social aspects of the society. Increased income is also an example of a direct measurable return, while the above example of a non-measurable benefit is also an example of an indirect return. The fourth dichotomy, positive-negative, distinguishes between positive returns of the system and negative results of the system. Negative results could be categorized as cost, but if cost is defined as the investment of resources and benefits as the returns, it is appropriate to classify negative results also as "benefits" and not as a cost.

Increased personal income and its effect on society is an example of a positive benefit. A list of examples of negative

benefits are associated with the student who becomes a system "push-out" (usually referred to as the "drop-out"). The decrease in potential earning and buying power, the resulting decrease in the income tax base, the potential increase in unemployment, welfare, crime, etc., the list of negative "benefits" associated with the "push-out" seems endless. Other negative "benefits", which are not measurable, result when the educational system hinders, instead of increases, the motivation and creativity of the student and when the system does not provide adequate guidance for students.

Costs of the system may be classified as: individual-society and measurable-non-measurable. The cost of operating the system, e.g., salaries, buildings, supplies, etc., are measurable society cost. The loss of property tax due to the use of space by the system and the interest paid on school bonds are also examples of measurable society cost. Money expended for personal needs, e.g., paper, pencils, books, student fees, etc., are measurable individual cost. An example of non-measurable society cost is the misuse of funds due to poor planning.

The number and the relationships of the variables which need to be considered in a complete classical cost-benefit analysis of an educational system presents an extremely difficult task.

Some of the cost-benefit studies of vocational education are listed in the references. These studies have provided data for planning and evaluating vocational education programs.

Is the classical cost-benefit analysis the only procedure for studying the investment of resources and the resulting benefits? The remainder of this section suggests another method for performing a cost analysis.

Since a strong intuitive case can be presented for the economic desirability of an education, the assumption is made that education is financially beneficial. A further assumption is that the characteristics of educated individuals who efficiently contribute to the society are related to the institutional objectives of the educational system. Therefore, the proposed cost analysis, referred to as Cost-Effectiveness Analysis, seeks to study the relationships between institutional objectives, instructional programs, measures of the achievement of objectives, and the cost of the system. (The parallel costs analysis, which seeks to ascertain and to compare specific economic benefits resulting from two alternative programs offered to the same population will be treated in the impact section of the evaluation. This analysis will require only estimates of benefits to the degree that benefit differences between vocational education and an alternative program catering to the same population of students can be observed,

and will be studied as part of the economic impact of vocational-technical education).

Analyses of the relationships between institutional objectives, instructional objectives, measures of the achievement of objectives, and cost could provide evaluators and planners with valuable data. Analyses of this magnitude are uncommon in traditional educational research. The above relationships do not have the attributes required of a two-variable, one-man project. Analyses of the above magnitude are not exceptional in the defense and space industries. Therefore, the adaption and application of the expertise of the system analyst to educational problems should provide the tools necessary to broaden the scope of the traditional researcher. One approach which may lead to the analysis of the relationships of the previously listed variables is presented in this paper.

The variables--institutional objectives, instructional objectives, measures of the achievement of these objectives, and cost--and their relationships define a system. To be more accurate, they define a subsystem of the larger educational system. To study this subsystem several assumptions are necessary.

(1) The school system has defined a set of institutional objectives.

(2) Instructional programs are based on sets of instructional objectives.

(3) Institutional and instructional objectives are stated in measurable terms.

(4) Instruments for measuring the achievement of objectives are available.

(5) Accounting and budget data is systematically maintained by the school system.

Cost is defined as the monetary resources allocated through a budgetary process for the operations of the school system. This definition excludes the cost associated with student time and resources, therefore, cost benefit analysis of the results of this approach to the analysis of the relationships of objectives and costs is limited.

There are basically three plans of operation in systems analysis, i.e., collecting data, categorizing and displaying data, and the analysis of data. If the analyst has determined precisely the analysis to be performed, the data to be collected can be defined and the format for categorizing and displaying the data can be designed. A plan for the order of operation on each of three levels can be designed and implemented. This approach is the traditional method of the educational researcher. However, there are cases where the precise analysis to be performed cannot be specified. A study of the data is needed before well-defined analysis can be initiated. The study is essentially collecting and attempting to categorize and display data. Categorization normally will suggest changes in

methods of collection as well as indicate additions and deletions to
the information being gathered. Therefore, the level of operation--
collection or categorization--is constantly changing. Once the
collection/categorization stage has been initially completed, then
analysis may be initiated. As analysis of the data progresses, it
is highly possible that the format of the data will have to be re-
vised and,in some cases, additional information will have to be
collected and categorized. The study, and more precisely, the col-
lection and categorization/display levels, is the approach outlined
in this paper for investigating the relationships between instruc-
tional objectives, institutional objectives, achievement of objec-
tives, and cost of instructional programs.

The matrix is one of the display formats used in systems
analysis. A matrix is a two dimensional array. It is used to display
values or characteristics which are associated with the pairing of
attributes of the two dimensions. Figure 1 is an example of a matrix
used to display cost data. The columns are the traditional activities
or functions of a school system. The rows are object areas or "line
items" of the system. The common method of recording cost for budget
and accounting purposes of a school system uses the activity/object
breakdown of cost. The sum of the row, "Salaries", would be the total
amount of money either allocated or expended for salaries in the sys-
tem. The sum of the column, "Administration", would be the total cost

Accounting Matrix

Activities

	Admini-stration	Super-vision	Instruc-tion	Bldg. & Grounds	
Salaries	$120,000				
Life Ins.					
Health Benefits					
Retirement					

Objects

Figure 1

of administering the system. The square in row "Health Benefits"
and in the column "Administration" would indicate the cost of health
benefits provided to administrative personnel.

The number of dimensions of an array may exceed the two of the
matrix. Sometimes it is beneficial to present data in a three dimen-
sional array. For example, it is possible to extend the previous array
of cost data by adding a dimension of "Instructional Programs".
Figure 2 is an illustration of this concept. Cost are related not
only to specific activities and objects, but they are also assigned
to specific programs. Many school systems have extended their
accounting and budgeting systems to include the program level. Hence,
administrators have readily available data pertaining to the cost of
programs. The advantages for planning and evaluation of instructional
programs should be obvious.

The variables being considered in this study may be displayed in
the matrix format. The first three matrices suggested for considera-
tion are:

(1) A matrix displaying a relationship between institutional
     and instructional objectives.

(2) A matrix displaying a relationship between institutional
     objectives and instructional programs.

(3) A matrix displaying a cost relationship between institu-
     tional objectives and instructional programs.
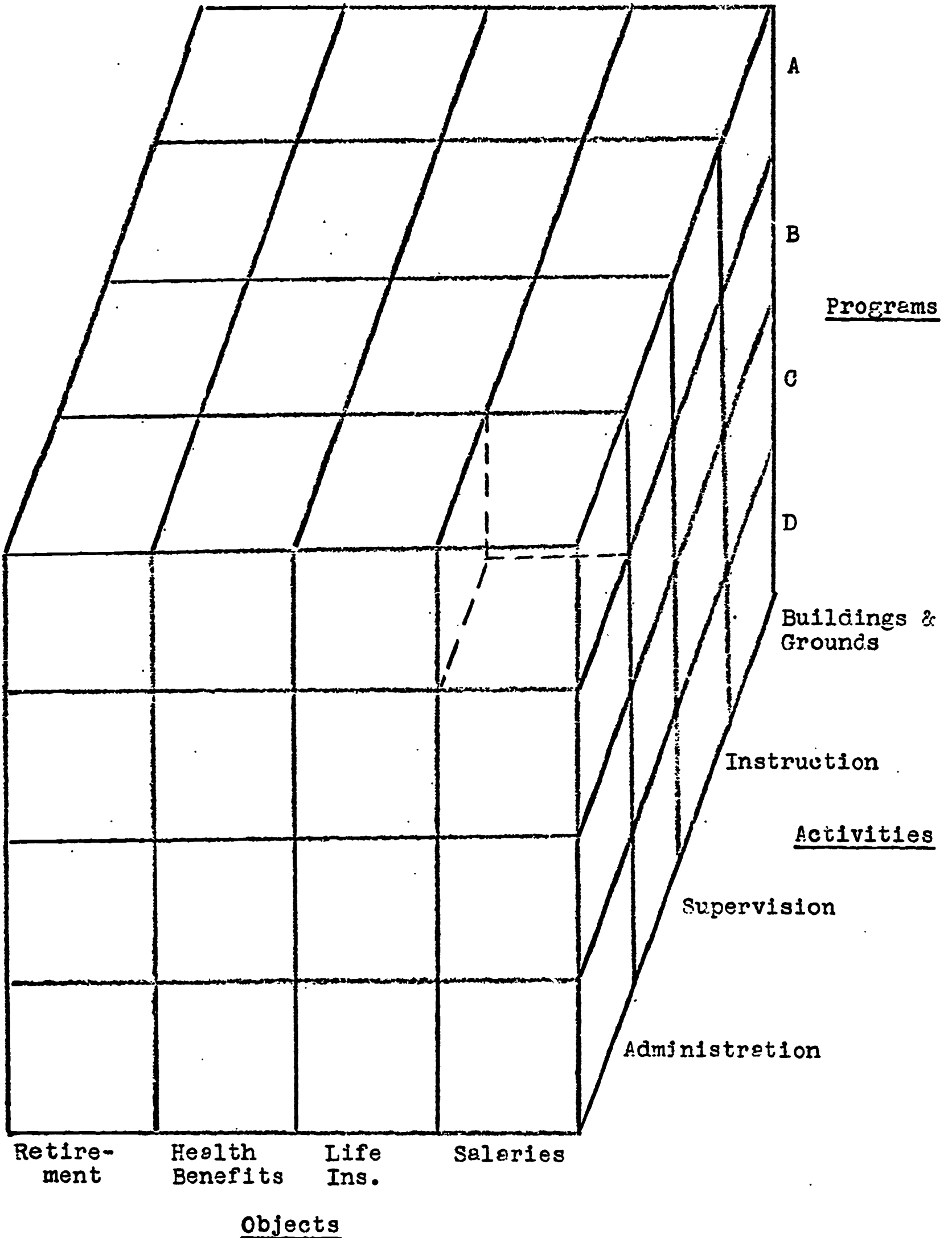
Three Dimensional Cost Display

Figure 2

Institutional objectives are achieved through instructional programs. Each program is based on a set of instructional objectives; therefore, it is possible to relate instructional and institutional objectives. This relationship is graphically illustrated in Figure 3. Each column of the matrix designates an institutional objective. Rows indicate instructional objectives of a specific program. The squares of the matrix contain one of the following values:

P--The instructional objective was planned to assist in
    achieving the institutional objective.

C--The instructional objective will contribute to the
    achievement of the institutional objective.

N--The instructional objective will have negligible or
    neutral effect in achieving the institutional objective.

Figure 1 illustrates a partially completed matrix for Program "A"-- a hypothetical instructional program. A similar matrix can be prepared for each of the instructional programs of an institution.

The set of matrices resulting from the above exercise may be combined in a matrix. Figure 4 illustrates a relationship between instructional programs and institutional objectives. The columns are the same as in the previous set of matrices. The rows designate instructional programs. The values of the squares are

Program "A"

Institutional Objectives

	a	b	c	d
1	P	P	C	N
2	P	P	C	N
3	P	C	C	N
4	P	P	C	N

Instructional

Objectives

Figure 3

subjectively determined from the data contained in the first set of matrices. A weight, from 0 to 10, is assigned to each block. If a column of a matrix describing the relationships between instructional objectives contains all P's, then a value of 10 would be appropriate. If a column contains all C's, a 4 may be appropriate. If a column contains all N's, a 0 would be appropriate. Using these three states as guides, values for each square can subjectively be determined. The weights assigned in the first row of Figure 4 are based on the data contained in Figure 3. Weights defined in the above manner are subjective. They are also relative to each other if the same subjective resoning is used in determining all weights displayed in the matrix. Care should be employed in interpreting the data presented in this matrix.

Cost factors can now be added to the display of data. The columns and rows of the next matrix, Figure 5, remain the same, but the values of the squares are determined from the cost assigned to instructional programs. The total cost of a program is proportioned over the institutional objectives according the the weights assigned in the previous matrix. This is accomplished by adding the weights of a row and dividing that sum into the total cost of the program. The resulting amount is then multiplied by each weight to determine the values to be assigned to the squares of the cost matrix. The sum of each row should equal the total cost of a program. For example, if the instructional objectives of Program "A" were only

Weight Matrix

Institutional Objectives

	a	b	c	d	
1	10	8	4	0	
2					
3					
4					

Programs

Figure 4

related to the institutional objectives for which data was presented in Figures 3 and 4, and if the total resources allocated for the program were $12,100, then the cost would be distributed as illustrated in Figure 5. To complete the cost matrix, sum each column. These sums are subjective indications of the budget resources allocated to the achievement of institutional objectives.

The next matrix proposed contains more data than the first three illustrations. One of the assumptions stated earlier was the presence of a measuring instrument to determine achievement of both instructional and institutional objectives. This matrix will present data pertaining to the success of students in meeting the stated objectives. Success is defined as a state where a student has either demonstrated he has reached a specified level of performance or that he has achieved a specified incremental increase in his level of performance on a specific objective. From this data the percentage of successful students of a given population can be determined. Several sets of percentage data can be derived. These are displayed in the following matrix, Figure 6.

The columns of the matrix designate institutional objectives. Associated with each column heading is the value indicating the successful percentage of the total student population. The rows indicate instructional objectives of a program. The associated value indicated the successful percentage of students enrolled in the pro-

Cost Matrix

Institutional Objectives

Figure 5

gram and who were tested on the instructional objective. The squares

contain two numbers. The first is an indication of the percentage of

students who were tested on the instructional objective and who

were successful on the institutional objective criterion. The second

number is similar to the first, but the student population is limited

to a subset of the previous population. The percentage indicates

successful students who are not enrolled in other programs which have

instructional objectives related, either in a planned or contributory

role, to the same institutional objective.

Another matrix which would be beneficial to a study of available

data should be similar to the three dimensional cost matrix presented

earlier. Instead of the third dimension being instructional programs,

it could be the instructional objectives of the instructional program.

For some cases the assignment of cost to instructional objectives may

be easily achieved, but the probability of having to subjectively

assign some cost factors is high. Therefore, a procedure similar to

methods previously described would have to be designed. Once data

is available in the above form it would be easy to generate a cost

matrix relating instructional objectives to institutional objectives.

This matrix would be similar to the matrix illustrated in Figure 5,

except the rows would designate instructional objectives instead of

instructional programs.

To complete the collection of data, it is suggested that a table

of instructional programs be compiled which would list the total cost

of the program and the level at which each program is funded. The

93

Achievement Matrix

Institutional Objectives

	a(73)	b(84)	c(91)	d(87)	
1(84)	77, 56	85, 83	92, 92	86, 84	
2(79)					
3(86)					
4(92)					

Instructional Objectives

Figure 6

level of funding could be determined from data generated by planners of instructional programs if requests for funds for programs were submitted in a format which would indicate different levels at which the program could operate.

The set of matrices and the table suggested for categorizing and displaying data would provide the analyst with a vast amount of data. The collection of data would be a time-consuming and sometimes difficult task. The study as defined would certainly require the application of data processing methods. Computer programs could be designed which would aid in the categorization of the data and would generate the displays in the suggested formats.

The third level of operation in systems analysis--the analysis of the data--is limited only by the imagination of the explorer. Working within the constraints of the validity of the subjective data, it would be possible to draw inferences and in some cases test hypotheses.

The purpose of this chapter has been to develop a framework for the analyses of cost-effectiveness relationships of instructional programs in vocational education. The completion of the proposed collection and display of data would provide the information necessary to design guidelines for a cost analysis of instructional programs.

CHAPTER SIX

PHASE IV:

EVALUATION OF VOCATIONAL-TECHNICAL EDUCATION

(Impact Study)

Jim C. Fortune

University of Massachusetts

The impact study will basically be concerned with ascertaining

the degree to which the Vocational-Technical Education effort is

meeting the needs of the Commonwealth of Massachusetts. Since the

needs are multi-dimensional in nature, several programs of study will

be directed toward the examination of impact. Since the impact

studies will generally deal with decisions of long-term importance,

rather than decisions producing immediate value for program modifi-

cation, these studies will be scheduled intermittently in the

development cycle of the evaluation and will be discussed below in

general terms.

The impact study has four components, namely: 1) a state-wide

survey of occupational needs, 2) a compilation of state-wide

statistics on enrollment and matriculation from vocational and

technical school, 3) several research samples designed to study

employment in areas of training, job satisfaction of graduates,

citizenship and social behavior of graduates, and aspirations of

potential graduates, and 4) a small sample longitudinal study of

program graduates done in a case-study manner.

The state-wide survey will include the development of a
questionnaire which will be sent to a random sample of major indus-
tries. This survey will seek employment and development information
which may prove helpful in program planning for vocational-technical
instruction. Additional knowledge on occupational needs in the state
will be sought from other state employment agencies whose data bases
offer potential enlightenment for program planning. Several sources,
including parents, students, educators, social workers, and econo-
mists will be surveyed to provide perspective to the industrial and
employment agency surveys.

The compilation of state-wide statistics on enrollment and com-
pletion of vocational-technical school programs for the Commonwealth
of Massachusetts is currently in progress. This report will further
demonstrate the total impact of the program for the state; it
eventually will include a catalog of vocational-technical offerings,
economic data on how vocational-technical graduates compare with
other educative program graduates, the data comparing the state
vocational-technical expenditures with those of other states, and
growth figures.

Two forms of analysis will be utilized to ascertain results in
terms of over-all vocational-technical programming goals. Random
samples of students will be followed after program completion and
measurements of their success on several variables will be made and

98

compared either to state norms or to non-equivalent control groups.
These include job-satisfaction, life-adjustment, personal asiprations,
and work records. A second form of evaluation will consist of direct
measurement of a multi-dimensional basis of the stated general goals
of vocational education. For example, citizenship training might be
measured in terms of several traits or variables, such as voting
registration and participation, civic club membership, credit records
and police arrests. In this instance, non-equivalent control groups
selected by matching age, race, and socio-economic status will form
the basis for comparison.

General program goals and impact employing a case-study method-
ology on a longitudinal basis will also be studied. A small random
sample of beginning students will be selected and biographic records
and diagnostic test files will be developed on each student. Teacher
and counselor reports on the sample study will be collected monthly.
For each subject, general standardized tests and indices which seem
pertinent to the goals under consideration will be administered and
recorded. Throughout the vocational-technical school careers and
several post school years, the sample subjects will then be studied.

CHAPTER SEVEN

EVALUATION, ACCREDITATION, AND

THE MASSACHUSETTS INFORMATION

FEEDBACK SYSTEM

A Commentary

James F. Baker
Assistant Commissioner

Research and Development Center
Massachusetts Department of Education

The purpose of this chapter is to discuss relationships between the accrediation process and the Massachusetts Information Feedback System. Within this presentation we will define accrediation as the endorsement of educational conditions, procedures, or programs by an authoritative group in light of published criteria or standards. We will narrowly define evaluation to mean both the description and judgment upon educational conditions, procedures, or outcomes. A precondition to evaluation is the process of measurement which is the quantification of conditions, procedures or outcomes, and is a process requirement before value judgments can be made. Measurement is a part of the description phase of the evaluation process.

The Massachusetts Information Feedback System has been thoroughly described in this publication. This system seeks to feed back information to both the local and state levels regarding achievement of pupils on specific behavorial objectives developed in the various vocational schools; it seeks to describe relationships between various teaching processes within the school and achievement (process-product evaluation); it seeks to describe the cost relationships between institutional objectives, instructional programs, and measures of achievement of objectives; and it seeks to describe the total impact

of vocational education on the Commonwealth. In short, the Massachusetts Information Feedback System attempts to provide a data base for a systems approach management to vocational education. The system does not attempt to restructure vocational education across the state and is extremely respectful of both local control, and existing state/ local relationships within vocational education. In the language of this chapter, the Massachusetts Information Feedback System objectively measures various aspects of the conditions, procedures and outcomes of vocational education and thereby describes vocational education in such a way that value judgments can be made frequently at all levels. Evaluation within the Massachusetts Information Feedback System is conceived of as an on-going process with information being fed back semi-annually.

Accreditation, the process of endorsing educational conditions, procedures, or programs, relies upon a description and measurement process for evaluation which is different than the procedures used by the Massachusetts Information Feedback System. Essentially, the the accreditation process relies upon the collective judgments of an authoritative peer group of a particular educational program, in light of published criteria or standards. These standards are usually modified to reflect the objectives of a local educational program.

Accreditation may involve a total school or a part of a total educational setting, as illustrated by an association's examination and accreditation of a program in business, data processing, engineering,

or teacher education. Accreditation activities are more often directed towards means rather than ends. Means, conditions, and procedures are observed and judged regarding their extensiveness and effectiveness. Judgments are made by self or visiting committee evaluation teams regarding the environment for learning, the activities contributing to learning, and the enthusiasm for the learning situations by the learners and those responsible for directing the learning.

The accreditation process may reflect activity by professionals over a significant period of time. For some school staffs one year's in-service activity may be directed toward the self-evaluation aspects of the accreditation study. Numerous committees may be developed to examine and report on specific education services (guidance, library, administration), or curricular fields (science, language, music).

Self-evaluation committees preparing for visitation by a peer group of professionals might spend 50 to 100 hours reviewing evaluation guides or check lists designed to provide some basis for judging the effect of the significant conditions and procedures operating within a school. Such periods of in-service activity are generally followed by two- or three-day visits by an official committee representing an accrediting association or state agency. Self-evaluation committees may involve all professional personnel in the school or program being studied or may be representative of the various facets of the institution. Visiting committees may be limited to two or three people

representing the accreditation group or may be expanded to about twenty-five percent of the staff of the school being studied.

Professional organizations reflect two points of view with reference to their accreditation responsibilities. Their concern is either the accreditation of the total institution or a specific program or curriculum within the institution. Associations accrediting specific programs are the National Council for Accreditation of Teacher Education (NCATE) or the Accrediting Commission for Business Schools. Such general associations as the Regional Association of Colleges and Secondary Schools, of which there are seven distributed geographically throughout the United States, represent groups which accredit schools at the secondary level or institutions of higher education.

As a rule, reports emanating from a visiting committee for accrediting associations are quite general in tone and usually identify recommendations and commendations regarding particular facets of the educational program for which a sub-committee has some responsibility. A trend has been exhibited during the last two decades for accrediting standards to become more general and flexible, reflecting standards which are more qualitative in nature. Such change in accrediting standards has been brought about because of the occasionally questioned validity of specific quantitative standards and with recognition of the inability of accrediting groups to develop single sets of criteria to meet all conditions or educational settings. As continued attention has been given to the accrediting process and as a very high percent-

age of schools and institutions have become accredited, standards for accreditation have become interpreted as minimal and educators interested in achieving goals beyond accrediting levels have sought a more objective description of programs.

The relationships between the Massachusetts Information Feedback System and the accreditation process are one of mutual support. It would seem reasonable to predict that, in the future, accrediting associations will rely heavily upon the information produced by a total educational information system. That is, the information produced by a total educational information system will constitute an objective data base by which an accrediting association can evaluate and endorse an education program. A total educational information system provides an empirical base for directed and controlled educational change. This information will become essential in the modification of existing vocational-technical educational programs. We would expect that accrediting agencies would consider the program modification process initiated by the local educational program in light of an information feedback system as a part of their total evaluation of a local educational program.

Both accreditation and an educational information feedback system are essential parts of the total educational process. An educational information feedback system seeks to provide information about the total educational process. It describes relationships between measured components of the total educational process, reports on the effective-

ness of education, and provides the data base for directed change and program modification. It is an essential planning tool. On the other hand, accreditation is a process of endorsing a particular school or program as being consistent with or better than established standards. Surely, the accreditation process will be more meaningful when more sophisticated techniques of measuring and describing vocational-technical educational programs become operative. The implementation of the Massachusetts Information Feedback System, or similar total educational information systems should provide a base to make accreditation a more meaningful process.

ADDENDUM

ADDENDUM

Progress in the Development of the System

Beginning in December, 1968, when the Research Initiation Team of
Vocational-Technical Education for the Commonwealth of Massachusetts
identified evaluation as a priority concern, the Research Coordinating
Unit proceeded to clarify the needs of evaluation in vocational educa-
tion. Several individual meetings with selected members of the Research
Initiation Team and with a consulting staff from the School of Education,
at the University of Massachusetts, led to the identification of needed
feedback on vocational-technical programming. Consequently, a systems
approach to the evaluation of vocational-technical education was devel-
oped, but due to the vastness of the project and the clear evidences of
much needed training prior to implementation of this system, it was
broken down into stages. Each stage, like the total system, was then
conceived to be a developing project and task analysis was utilized to
decide optimal starting points.

Since the individual program assessment was designated as the most
needed feedback, since the provisions designed to allow local autonomy
in curriculum control in individual program assessment created the
greatest number of logistic problems, since data from individual pro-
grams was needed in almost every phase of the evaluation system and
since behavorial objectives were necessary for individual program
assessment, the total system was started by introducing a six school
pilot program at the Annual Winter Conference of Vocational-Technical
Education at Auburn, Massachusetts, February, 1969. The six pilot
schools will be operational during the 1969-70 school year.

A workshop on objective writing and a follow-up conference were
held in April, 1969, at Northampton, Massachusetts. Each of

the six pilot schools designated three facilitators who attended the workshop to prepare them for the coordination roles of developing behavioral objectives for the programs at their schools. Currently, these men are developing the objectives and the Research Coordinating Unit is planning the test development procedures.

The conference schedule at Northampton included the following sessions:

What is Evaluation?
What are Behavioral Objectives?
Occupational Analysis and Behavioral Objectives
Small Group Work Session--Objective Writing
Introduction to Taxonomies as a System for
    Classifying Behavioral Objectives
Review: Task Analysis, Taxonomies, and Behavioral
    Objectives
Review: Total Evaluation Information Feedback System
Introduction to Simulated Objective Writing
Conference Participants Develop Objectives Under
    Simulated Conditions
Process-Product Evaluation
Simulated Objective Writing
The Relationships Between Evaluation for Accreditation
    and Evaluation as an Information Feedback System: a
    panel discussion

The conference participants and staff were:

Don Glazier	Animal Science	Essex Agr.& Tech. Inst.	Hathorne
Dick Churchill	Plant Science	Essex Agr.& Tech. Inst.	Hathorne
Dave Twonley	Forest Resources	Essex Agr.& Tech. Inst.	Hathorne
Russell J. Booth	Machine Shop	Diman Regional	Fall River
Edward Peckham	Animal Science	Smith's Agr.& Voc.	Northampton
Frank Shore	Basic Electronics	Haverhill Trade	Haverhill
Gerard B. Lachance	Electronics	Diman Regional	Fall River
Henri J. Pare	House & Mill	Diman Regional	Fall River
Jim Boland	Electrical	Haverhill Trade	Haverhill
Edward Vandoloski	Electrical	Smith's Agr.& Voc.	Northampton
John Filipek	Auto	Smith's Agr.& Voc.	Northampton
Robert P. Nelson	Technical Coor.	Blue Hills Reg.	Canton
Arthur Vuilleumier	Elect. Tech.	Blue Hills Reg.	Canton
David Malone	Math & Science	Blue Hills Reg.	Canton
Michael J. Mango	Auto Mechanic	Chicopee Compr.	Chicopee
Quinto Cimma	Welding	Chicopee Compr.	Chicopee
Thaddeus Ossolinski Jr.	Drafting	Chicopee Compr.	Chicopee
Normal J. Campana	Sheet Metal	Haverhill Trade	Haverhill

Staff

Jim C. Fortune
School of Education
University of Massachusetts

Robert Schweiker
Measurement and Evaluation Department
Ontario Institute for Studies in Education

Ray Johnson
School of Education
University of Massachusetts

Mary E. Green
Massachusetts RC Unit
Massachusetts State Department of Education

Joanne Szarlan
Massachusetts RC Unit
Massachusetts State Department of Education

William G. Conroy, Jr.
Massachusetts RC Unit
Massachusetts State Department of Education

A word of special thanks is due to Walter Markham, State

Director of Vocational Education; James Baker, Assistant Commissioner,

Department of Education; and the six directors of the pilot schools:

William A. Dwyer, Director
Blue Hills Regional Vocational High School

John Connolly, Director
Haverhill Trade School

B. Stanley Dowgert, Director
Northampton-Smith Agricultural School

Alfred R. Rios, Director
Chicopee Vocational High School

James F. Gallant, Director
Essex Agricultural and Technical Institute

John P. Harrington, Director
Greater Fall River Diman Regional Technical School

The Research Initiation Team is composed of:

John P. Harrington, Director
Greater Fall River Diman
   Regional-Technical School
575 Prospect Street
Fall River, Massachusetts

William A. Dwyer, Director
Blue Hills Regional Vocational
   High School
Canton, Massachusetts

Michael Gonzalez, Director
Westfield Trade School
Westfield, Massachusetts

Donald E. Graves, Director
Southeastern Vocational
   Technical High School
Brockton, Massachusetts

Paul Sullivan, Director
Blackstone Valley Vocational
   Technical High School
Upton, Massachusetts

Ruth Shea, Director
Henry O. Peabody School
   for Girls
Norwood, Massachusetts

Ann McDonald
Assistant Superintendent
   of Schools
Brookline, Massachusetts

Included are the following State Department personnel:

Walter J. Markham, Director of Vocational Education
Grace L. Nangle, R.N., Senior Supervisor, Health Occupations
   Education
Helen Jean Marks, Supervisor, Home Economics
Philip A. Haight, Supervisor, Agriculture
John P. Manning, Supervisor, Adult Distributive Education
Robert L. Manning, Senior Supervisor, Business and Office Education
John Fitzgerald, Senior Supervisor
John P. Morine, Senior Supervisor, Vocational Guidance

BIBLIOGRAPHY

BIBLIOGRAPHY

Administrative System for the Model School Division of the
    District of Columbia Public Schools, prepared for MSD
    DC Public Schools, General Learning Corp., Washington,
    D.C., 1968.

Ammerman, Harry L. and William H. Melching. The Derivation,
    Analysis, and Classification of Instructional Objectives.
    Human Resources Research Office, The George Washington
    University. U.S. Army contract DA 44-188-ARO-2,
    Technical Report 66-4, May, 1966.

Anthony, R.N. Planning and Control Systems: A Framework
    for Analysis. Division of Research, Graduate School
    of Business Administration, Harvard University,
    Boston, 1965.

Bloom, Benjamin S. "Learning for Mastery". UCLA Evaluation
    Comment. Center for the Study of Evaluation of
    Instructional Programs, The University of California,
    May, 1968, Vol. 1, No. 2.

Breuhaus, B.A. "Accreditation Enhances Quality". NEA
    Journal, 49:52-53, April, 1960.

Corazzini, A.J. Vocational Education, A Study of Benefits
    and Costs (A Case Study of Worcester, Mass.).
    Princeton University, August, 1966.

Cost Study of Educational Media Systems and Their Equipment
    Component, Volume 1, Guidelines for Determining Costs
    of Media Systems. Prepared for U.S. Department of
    Health, Education, and Welfare, Office of Education,
    Bureau of Research. General Learning Corp., Washington,
    D.C., May, 1968.

Cronbach, Lee J. "Course Improvement Through Evaluation".
    A working paper unpublished in present form. University
    of Illinois, 1962.

Davis, R.G. Planning Human Resource Development. Rand
    McNally and Co., Chicago, Illinois, 1966.

Deferrari, Roy J. Self-Evaluation and Accreditation in
    Higher Education. Washington, The Catholic University
    of America Press, 1959.

Dickey, F.G. "Accreditation and Excellence". AACTE Yearbook,
    128-134, 1962.

Educational Evaluations: New Roles, New Means. NSSE Yearbook,
    Part II. University of Chicago Press, 1969.

Eninger, M.V. The Process and Product of T and I High
    School Level Vocational Education in the United States --
    The Process Variables. Educational Systems Research
    Institute, American Institutes for Research, Pittsburgh,
    Pennsylvania, 1964.

Guba, Egon G. "Development, Diffusion, and Evaluation".
    A paper presented to the Conference on Knowledge
    Production and Utilization in Educational Administration:
    Role Emergence and Reorganization. The National
    Institute for the Study of Educational Change,
    Bloomington, Indiana, 1968.

Harris, C.W. Problems in Measuring Change. University of
    Wisconsin Press, 1963.

Lindvall, C.M., editor. Defining Educational Objectives.
    Report of the Regional Commission on Educational
    Coordination and the Learning Research and Development
    Center, Pittsburgh, University of Pittsburgh Press, 1964.

Mager, Robert. Preparing Instructional Objectives. San
    Francisco, Fearon Publishing Co., 1962.

Morris, Alvin E. A Critical Analysis of Types of Quantitative
    Data Requested by the Professional Teacher Education
    Accrediting Agency and the Six Regional Accrediting
    Agencies. Detroit, Wayne University Press, 1959.

Morse, P.M., editor. Operations Research for Public
    Systems. The M.I.T. Press, Cambridge, 1967.

NCTEPS. Milestones in Teacher Education and Professional
    Standards. Washington, National Education Association,
    1964.

Pace, C. Robert. "Evaluation Prespectives: '68". A
    working paper unpublished in present form. University
    of California, 1968.

Payne, David A. and Robert F. McMorris, editors. Educational
    and Psychological Measurement. Blaisdell Publishing
    Co., 1967.

Selden, William K. Accreditation, A Struggle Over Standards
    in Higher Education. New York, Harper and Brothers, 1960.

Selden, William K. "The Three Basic Principles of Accrediting".
    School and Society, 87:308-310, June 20, 1959.

Stake, Robert E. "The Countenance of Educational Evaluation".
    Teachers College Record. Vol. 68, No. 7, April, 1967.

Stufflebeam, Daniel L. "Evaluation as Enlightment for
    Decision-Making". An Address to the Association of
    Supervision and Curriculum Development, Sarasota,
    Florida, January, 1968. Evaluation Center at Ohio
    State University, Columbus.

Zook, George F. and M.E. Haggerty. The Evaluation of Higher
    Institutions. Chicago, University of Chicago Press, 1936.