DOCUMENT RESUME

ED 046 968 TM 000 127

AUTHOR Petry, John R., Ed.
TITLE Proceedings of the Evaluation Workshop Retreat.
INSTITUTION Memphis State Univ., Tenn.
PUB DATE Jul 69
NOTE 124p.; Proceedings of the Evaluation Workshop Retreat, Burns, Tennessee, July 1969

EDRS PRICE EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS Behavioral Objectives, Data Collection, Decision Making, Educational Improvement, Educational Objectives, Educational Planning, *Educational Programs, Evaluation, Evaluation Criteria, *Evaluation Techniques, Management, Problem Solving, *Program Evaluation, Program Planning, *Research Methodology, Statistical Analysis, *Systems Analysis, Systems Development, Tests, Workshops
IDENTIFIERS *Elementary Secondary Education Act Title III, Tennessee

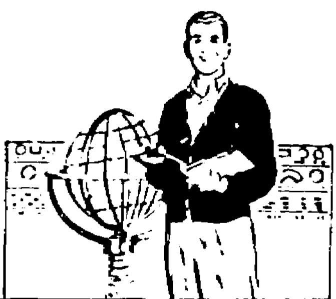
ABSTRACT

The Design for Tennessee Assessment and Evaluation of Title III, Elementary and Secondary Education Act, was developed to provide a tool for educators to use in evaluating projects and programs in Tennessee schools. An intensive workshop on evaluation was held using the Tennessee Evaluation Design as the synthesizing structure. The presentations began with an overview of evaluation and a discussion of the design itself. Other presentations included "Planning Networks and Program Management," "State Evaluation, a Team Effort," "Writing Educational Objectives," "An Organizational Structure of Variables Affecting Educational Programs," "Measurement Instruments," "Application of Analytic Techniques," and "Data Collection Techniques." A glossary is included. (PR)

Proceedings Of The Evaluation Workshop Retreat

July 28-30, 1969

Montgomery Bell State Park
Burns, Tennessee



WORKSHOP DIRECTOR: Fred K. Bellott

EDITOR: John R. Petry

Memphis State University
College of Education
Bureau of Educational Research and Services
Memphis, Tennessee

Proceedings Of The Evaluation Workshop Retreat

July 28-30, 1969

Montgomery Bell State Park
Burns, Tennessee

Editor:
John R. Petry

Workshop Director:
Fred K. Bellott

INTRODUCTION

The *Proceedings of the Evaluation Workshop Retreat* is the third of a series of publications dealing with Title III, E.S.E.A. programs, the first two being *Design for Tennessee Assessment and Evaluation of Title III, E.S.E.A.* and the abstract of that document. This document records the formal speeches of the workshop personnel but not the comments from the discussions that followed. Each speech is retained in its program order in an attempt to recapture the major impact of the programming procedure. All articles have been edited by the authors previous to the inclusion in this work. Only those by Desmond L. Cook differ from his presentation; however, his three articles form the basis for his conference remarks.

The majority of the ninety-three participants were nominated by the superintendents of their respective school districts to attend the workshop because of their involvement in Title III, E.S.E.A. programs. They represented forty-two school districts and thirty-six counties from the three grand divisions of Tennessee. Others present were employees of the Tennessee State Department of Education, consultants, and faculty and staff members of Memphis State University.

The beginning of the workshop established the purpose of the meeting and set the climate for maximum acceptance of the events to follow. Subsequent speakers emphasized instructional presentations on evaluation design components. These definitive statements are fundamental to an understanding that evaluation is a vital aspect of project success. Ample opportunity was given in group meetings and in leisure hours for amplifying ideas, resolving semantic differences, illustrating methods of approaching problem-solving, and sharing successes and failures in project management.

It is hoped that this book will be more than a remembrance of three days at Montgomery Bell State Park. It is intended that it be an additional resource for attacking problems that confront us in all aspects of educational management, whether our involvement be direct or indirect, whether our role be substantial or minor.

From this document, a series of reprints will be forthcoming, primarily for selected target populations who will use them in preference to the larger, more comprehensive volume, in conferences with a limited scope.

John Petry, Editor

PREFACE

The publication in the Spring of 1969 of *Design for Tennessee Assessment and Evaluation of Title III, E.S.E.A.* was the result of a developmental research effort to provide a tool for educators to use in evaluating projects and programs in Tennessee schools. Several regionalized meetings were held across the state to give local school personnel an initial exposure to this evaluative tool.

It was evident that there was a need for further consideration beyond these brief meetings. Therefore, the Evaluation Workshop Retreat was arranged to better meet this need. The workshop was an intensive three-day institute on evaluation using the Tennessee Evaluation Design as the synthesizing structure.

The workshop was organized around the use of consultants for presentation of component topics during general sessions. These were followed by small-group work sessions under the leadership of discussion group leaders.

Evaluation of the effectiveness of the workshop showed significant positive results in a large majority of the components measured. The transcriptions included in this proceedings document can be useful to project personnel who have evaluative responsibilities but did not attend the workshop. They can also serve as review and reference sources for readers who desire a means of follow-up and reinforcement.

Fred K. Bellott, Director
Evaluation Workshop Retreat

SYNOPSIS OF WORKSHOP PROCEEDINGS

This document includes the transcriptions of presentations given by the consultants at the Title III Evaluation Workshop at Montgomery Bell State Park. What may not be clear to the reader at this time is that these presentations served as focal points for small, group discussions which took much of the time in the workshop and were one of the most important parts of the workshop. In these small groups, educators from across the state discussed with consultants the issues presented in the speeches. These small group sessions were geared to the actual, "at home" problems of evaluation. Of primary concern for the small groups were the writing of behavioral objectives and the systematic collection of data.

Dr. Colmey's overview of evaluation and Mr. Jones' plea for a team effort in evaluation set the stage for discussions of these aspects of evaluation. With the move at the national level toward multi-dimensional criterion measures, more frequent use of observation techniques, implementation of conclusion-oriented evaluation, and above all, process evaluation, the opening address was particularly appropriate.

The need to consider evaluations from the conception of the project through the final evaluation was stressed by Dr. Bellott in his presentation of the Tennessee Evaluation Design. For those who were so inclined, the comparative relationship of the Tennessee Evaluation Design and the EPIC Model could be recognized after Dr. Cornell's introductory remarks. The major points stressed in both presentations were the need for preliminary planning in the determination of goals, the need for communication through objectives during the project, and the need for varied types of observations during the operational phase of the project.

Dr. Cook offered help in the administrative problems of project management and Dr. Cornell stressed the writing of behavioral objectives from two viewpoints. Both, however, pointed to the need for the educator, as a strategist, to accurately define what types of behaviors will be required of the learner in specified situations. Much time was spent in the small discussion groups reacting to the presentations about the logic and technique of writing behavioral objectives. From the reactions of the participants and the percentage of time spent on the subject, the writing of objectives must be ranked as one of the most important topics of the workshop.

The other speeches by Drs. Nowell, Bowman, and Williams dealt with types of data, methods of data collection, and statistical methods for analyzing data. These topics, of a more technical nature, were also considered in the small groups because they relate to the patterns and uses of behavioral objectives and evaluation. The relationships among the objectives, measurement techniques used, and selection of appropriate analyses were stressed, providing appropriate points on which to end the workshop.

Terry Bond, Research Specialist
Memphis City Schools

CONTENTS

LIST OF FIGURES

7

1
EVALUATION: AN OVERVIEW

by
James W. Colmey

EVALUATION: AN OVERVIEW

James W. Colmey, *Director*
Bureau of Educational Research and Services
Memphis State University

Some general remarks may give you a little better understanding of why we are here and what this Evaluation Workshop should accomplish. What are we here to talk about? We know it is "evaluation,' but let us probe a little further to see what this "monster" called "evaluation" really is.

There is a tendency to take old terms that we have used and to change them into new terms or to redefine them. In reality, we are really working with the same kinds of things that we have done through the years but with considerable improvements and modifications.

We are talking about descriptive data. How do you describe the program that you are conducting? That is essentially what it is all about. How do we describe something—we observe, collect, write, define, measure, manipulate, present, and analyze. Obviously, we do some other things, but essentially these just plain words say what it is all about.

Observation can be a very simple thing. For example, I am observing you and you are observing me. But observation through rigid logic, or disciplined by controls and cross references, becomes something a great deal more complex. In reality, we are doing the obvious things which we have done for years, but we are becoming a great deal more accurate and objective.

The men that are on your program will make presentations that will extend each of these words that I have casually mentioned, and I am quite sure that you will experience some anxiety, confusion and concern about what they have to say. First, some words will be used differently from the way you use them traditionally. Some words you may not have used before. In addition to this, there are some techniques that you may not have used. This does not mean that these techniques are entirely different from what you have been doing—they are just modifications of what you have been doing in your everyday activities as teacher, supervisor, or superintendent through the years.

In the next decade, we are going to depend upon specialists (such as the men on your program) who have worked hard enough on particular techniques that they can present them with clarity to large groups (such as this one). In today's world, new knowledge must be disseminated more rapidly than in previous decades. What you learn in these three days, you are going to have to turn around and teach to somebody else. Changes can and must take place within five or ten years that formerly took fifty years.

When talking about new systems, approaches, modifications, or terminology, we have a tendency to forget the importance of existing documents. Do not take weak substitutes for important documents that have been available to you through local school systems and the Tennessee State

Department of Education in the past decade, such as Tennessee's Annual Statistical Report. These documents are a vital part of your evaluating system. These are basic to the evaluative formats that we are trying to project and evolve into a more advanced form.

Who does the evaluating? There is nothing exclusive about an *evaluative process* and Title III, E.S.E.A. Because there has been more emphasis on evaluation for Title III projects, there sometimes is a unique identity for evaluation in these projects. Efforts to improve instruction are going to increase the need for local, state, regional, and national evaluations in all educational programs. There is no monopoly on who evaluates or what is evaluated. However, each of these different segments of education will require unique adaptations. You will also find that there will be a greater need to try to coordinate the activities in evaluation at the local level with the kinds of activities that are going on at the state level. At the local level you are going to be looking for some different things than at the state level, but you must have some common ground in order to have both of these needs fulfilled. So, in plain language, we will all be evaluating in different ways.

Why do we evaluate? There is only one *real good* reason. There are other reasons that are not so important. *We need to have more meaningful information to improve instructional programs and to improve education for boys and girls in the state.* Other things must in some way be related or else they are rather meaningless.

In terms of objective evaluation with real meaningful purposes, efforts should not be to determine if this is "good" or "bad"—if it is "bad" throw it out; if it is "good" keep it. It is not that simple. It is not a meritorious kind of thing where "you get a blue ribbon," "you get a red ribbon," "you get a white ribbon." Hopefully, we have gone past that; hopefully, we can deal more objectively and meet what the public is demanding at this point.

Our general citizenry, legislators, teachers, and children are demanding improved educational programs. The main reason is that teachers today are better educated than they used to be and so are our citizens, legislators and children. We have had traditional jokes about incompetent state legislators—and certainly we knew cases where they have been true and during some periods in history where collectively it may have been pretty close to true. But I assure you that we have legislators in Nashville today who know just as much about education as most of us. They are "hardheaded" and they want answers because they know that, when they go back and talk to people in their community, they are going to have to talk to people who are more intelligent and are looking more objectively for the kind of answers that they want than used to be the case. It is just that simple.

I used to think, when I was a young school administrator twenty years ago, that if I could just get intelligent, well-educated teachers, I would not have any more problems. That was one of the biggest mistakes I have made. The more intelligent and educated your teachers and students, the more insight and questions they have. It does not take them long to call your attention to problems and opportunities. Therefore, you must be more straightforward than used to be the case. Within the next ten years, this will be a much more important consideration in Tennessee than in the past ten years. Fact gives merit to the kind of conference that we hope to have in the next few days.

Next question, how do we go about this process? The men on your program are going to answer that question, but I can give you a few generalities that will be the background for their remarks. First, they will be talking about evaluating as a never ending series of assessments, measurements, and analyses. Participation is all going to become more complex, useful, and objective. There will be more frustrations in this process, but, presumably, better decisions can be made because of it.

Some educators say that we have an infinite number of variables to deal with; therefore, research discipline does not mean anything if you try to be objective. Some will carry clinical research concepts to the point of having no utility in conclusions. Either extreme is unrealistic. The fact that educational research and evaluation cannot be as precise as physical science does not mean that education can not improve itself through an orderly process of observation, measurement, and analysis. We have got to believe that.

A weak spot in educational evaluation is that many components are appropriately accomplished but frequently left unrelated. We have a tendency to work in isolation. One group does testing, another group writes objectives, another group makes decisions. We do not always put these efforts together, and we do not always get the benefits from a strong contribution. In the next year or two, your state and local evaluation activities should be more closely coordinated. Tennessee has had a better state-local relationship than I have observed around the country. I think it is fair to say that in the last few years this relationship has improved. We have had more opportunities to work together. It is appropriate that universities participate and contribute, and it is appropriate that other agencies of the state, such as welfare, begin to work with state and local educators.

In the next three days we will be talking about tools for management, some improved tools that have been available to us and some new tools. We are trying to demonstrate new dimensions. The only reason for doing this, if we are honest and sincere, is to improve education for boys and girls. If you keep these things in mind and if you recognize that the presenters of material at this workshop are specialists who are going to bring to you specific kinds of information, you will benefit more from this workshop. The presenters are searching for ways to help us, and what they say will have more meaning to you, if you think of their remarks within the context of the comments I have made this morning.

2
THE TENNESSEE EVALUATION DESIGN

by
Fred K. Bellott

THE TENNESSEE EVALUATION DESIGN

Fred K. Bellott, *Associate Director*
Bureau of Educational Research and Services
Memphis State University

Thank you, Dr. Colmey. Much of our work is carried out in a team effort and members of the team cannot work in isolation. If we did try this, I am afraid that we would not get very far. The planning and conducting of this workshop was very much a team effort, not only in the Bureau but also between the Bureau and the Tennessee State Department of Education and all consultant members of the workshop staff.

I would like to make a few comments in reference to the pre-test that we were working with earlier this morning. First, I should explain that this was made up from questions submitted by each person who will be making a presentation during these three days. From the something like ninety questions that were submitted, the pre- and post-tests were selectively compiled. This is the primary reason that the tests have a varied format and a varied type of response. We tried not to constrain those people who were supplying the questions in using a specified format. They were given complete freedom in formulating their test items. Another reason was that we wanted to demonstrate several formats and the use of different kinds of measures. You may have noticed that the last area in the test dealt almost exclusively with affective behavior while the others were objectively oriented to cognitive behaviors. The last section also illustrates a semantic differential instrument which utilized a Likert-type scale as the means of response.

There will be a post-test using a similar instrument on Wednesday. It is from the comparison of the two that we would be able to draw some conclusions about our success in this three-day workshop. Again, with reference to the pre-post design, we are evaluating the staff, not you, the participants.

I would like to build upon several things that Dr. Colmey has said in terms of a frame of reference for the Tennessee Evaluation Design. First of all, I think that we should recognize the significance that is attached to the name of the model we will be working with in this workshop--The Tennessee Evaluation Design. It has been submitted to, and some affirmative endorsement has been received from, the Tennessee State Department of Education as a design that is appropriate for use throughout the state with most projects that are utilizing evaluation procedures. The design itself relates quite closely to several others with which you may be familiar. The scientific method provides one such comparison. Jack, I remember your mentioning this earlier in our discussions: there is a close parallel here between the step procedures found in the evaluation design and what we have recognized as the scientific method. If you are oriented to data processing and some of the procedures utilized, you might think of this as more closely related to the systems approach. Again, in the step by step sequence it is quite closely related.

The Tennessee Evaluation Design has four phases. We have called them Status evaluation, Planning evaluation, Operational evaluation, and Final evaluation. They are similar in nature. Earlier this year, the State Department, along with the State Testing Bureau, and the Bureau of Educational Research and Services, disseminated an instrument to gather information for a "state-wide assessment of educational needs." The assessment instrument went to a sampling population of about 4,800 randomly selected teachers, plus all principals, counselors, and superintendents in the state. This was the first move toward Status evaluation for the state on a stratified basis. Status evaluation demands the definition of the operational context to determine what it is we are working with, and where we are. It should serve somewhat the same purpose as our pre-test this morning—to establish the pre-existing conditions.

From this we move toward the identification of needs. If we had in fact made a proper type of assessment today, it should have been reflected in our program. The program would presumably be designed to meet the needs which had been identified.

Using the data collected in the assessment and also utilizing other existing data, such as in the annual statistical reports and other kinds of resources for what I call "base-line" data, we should identify what needs exist. These base-line data (Dr. Colmey earlier introduced this element), we feel, are very important. There are many areas wherein we do not have base-line data currently existing. Where this situation does exist we need to build into our Status evaluation the collection of base-line data; but, where these data already exist, there is not a reason in the world why we should not use them. Sometimes, I am afraid that we frequently go through the exercise of rediscovering the wheel simply because we do not go to the trouble to find out what has already been done. If we are asking twenty items that were just asked a month ago, or six months ago, or last year, the data is still valid data if the situation has not changed since that time. The description of the context remains the same and, where these situations exist, I certainly encourage the use of existing base-line data without going back and getting it again. Again, our problem in Status evaluation is to determine where we are and what our problems are.

We move from that into Phase II, or Planning evaluation. Once we have identified needs, what are we going to do about them? What do we do? It is at this point that our objectives begin to emerge. In order to deal with the goals that we would define here, we certainly need to look at several ways of attaining these goals and how we would define the objectives. Then we have to consider the alternatives that are available to us, not only in terms of a general design but also in terms of a specific strategy. Many times we belatedly find ourselves in a position of having to make evaluation after the fact. For example, someone who has been operating a project for two years suddenly realizes, "we have got to evaluate." What does one do in a week's time to get an evaluation made?

We feel that this is an inappropriate approach to the task of evaluation. Evaluation ought not be imposed after the fact; it ought to be included at the beginning of a program or project. At the time the objectives are established, we should have a [illegible] clear idea about what other activities we are going to evaluate and ho[illegible] [illegible] will be done. We may want to use a pre-post design. If the decision is made after the fact, it may be that the only thing one can do is to take a post-test. But if we do not know where we were to start with, how do we measure gain? This is not to say that the use of the post-test design is necessarily inappropriate. There are other applications of design that can yield much more information to us if we have the use of foresight instead of hindsight.

Planning evaluation includes thinking through these types of considerations. We need to determine the strategy of problem solution and the design that is to be used in operational evaluation—the means of measurement. Are we going to use standardized tests as the only type of objective measurement? Will we utilize subjective measurements taken in process? These determinations should be made before the project gets "down to the wire."

Let us move now to Phase III, Operational evaluation. We have our project underway and we are evaluating (I use the term as a description of our continuous evaluative activities). We are not only aware of what we are going to do, but we are in the process of doing it—because Operational evaluation is a process, a continual process. It is not a "one-shot" approach. From the time your project becomes operational, you have several tasks in developing and/or selecting the instruments that will be used for measurement and periodic observations.

Dr. Colmey dwelt on the term "observation" for a while and I would like to expand that a little. We like to think of the term "observation" in a broadened context, not as just looking at something but as a *recorded description*. I can look at Terry (Dr. Cornell) and, as I observe him, I may or may not go through an evaluative process subjectively in my mind. I consider that the three letters I received from him this past month are sources for observations. I hope to have additional correspondence and these will enable additional observations of him to be made. These examples are personal observations which Dr. Cornell and I have made recently. If we were asked to evaluate something to which these observations were relevant, they would be useful.

So let us use the term broadly, not just limit its application to one kind of operation. An observation may be in the form of the pre-test that we had this morning; it may be interpreted simply as our looking at one another. But an observation is an important element of Phase III, Operational evaluation.

From the observations that we have made, presumably (if this is to have real meaning to us) we will have to have some type of analysis. We will have a record of what has transpired during this project which will include all of the observations made. But the analysis itself is another step in Phase III. Analysis means to break down and look at the operation microscopically, to separate it into its parts. If you are not going to separate the parts, you simply get a total spectrum of the whole and cannot see it analytically.

After the analysis of the data, though, we must synthesize the information and consider it in the context from which it came. In the same way that the references were made earlier to "not working in isolation," we have to put these data together as emphasized so that they give us meaningful information about our total effort, our total project. In this way, we can have a feedback of information from the analyses for the purpose of decision-making. If we have analyzed scores, for example, on performance in a reading project and have found that we are not approaching our objective, then we feed this information back into the cycle. We can make modifications in the project at that time. We would do the same thing with our evaluation in any project, if we find that we were not getting the kinds of measurements that we had anticipated. On our recycle we have the advantage of the feedback.

This is a matter of communication, and so often this is where our own organizations begin to come apart at the seams. We do not practice what we preach in the use of feedback. The administrator does not wait for the feedback of information before he makes a decision; or, the testing consultant, after he has received his data and analyzed it, lets it stay in the drawer. These kinds of blockages continually harass and prevent us from getting the whole evaluation process into operation in a meaningful manner.

Phase IV is Final evaluation. The label used here, Final evaluation, was purposefully chosen. We are looking now at the whole project. We have a series of observations because we have monitored the project in process. Now it is time to look at the total effort and see what it all means. We were collecting data, then, for the total p oject. We will analyze the data concerning our pre-post status, or change measures however they were determined. Then we can interpret these data *in terms of the criteria that we had previously established.* These are the criteria that were set forth in Phase II in our objectives. They may have been modified since being written down originally, but they represent the kind of achievement that we have expected. These criteria (that have been determined already) should be used in interpreting the data and arriving at a point of decision. From these we should have a recommendation about the next cycle or to the administrator who has to make a program decision.

It is my feeling that information from evaluation ought to be conclusion-oriented. It ought to enable us to determine an appropriate course of action —whether what we have done was a "good way of doing" or whether the ways in which we previously attacked the problem were better, measured by how much success we attained in it. It also ought to enable the administrator to make a decision that would justify our support and our money or other resources. Is this the best alternative? If the evaluation is not decision-oriented, in my opinion, it has questionable value. This has application to the local level as well as state and national levels. I think that, if we cannot provide evaluative information to the State Department by which they can make program decisions, we have no right to expect them to decide in favor of our program. The same kind of application, of course, applies to your local school boards.

The use of the information that comes from this evaluative process, I think, is the most important element in the whole design. If you do not use it, you should not have gone to the trouble to gather it. It has to become meaningful by affecting the program development and the decisions that are made for support of them, and for the development of further programs or the broader scale of implementation.

I have tried to sketch through the design and its four phases and the components with which we will be working and will deal with in detail in forthcoming sessions. During the afternoon session, Dr. Desmond Cook will be with us. Dr. Cook has worked with planning in a particular frame of reference, that of management. Some of his techniques have been utilized in our own projects in development.

3
PLANNING NETWORKS AND PROGRAM MANAGEMENT

by
Desmond L. Cook

The following articles formed the basis for the presentation by Dr. Cook.

BETTER PROJECT PLANNING AND CONTROL THROUGH THE USE OF SYSTEM ANALYSIS AND MANAGEMENT TECHNIQUES[1]

Desmond L. Cook
Educational Program Management Center
The Ohio State University

INTRODUCTION

It was my good fortune this August to have been able to attend a five-day course at UCLA focusing upon a critical appraisal or state-of-the-art review of management information systems as they are now developing in the area of our society most commonly referred to as the military-industrial complex. The forces of fate, or just some simple clerical operation, resulted in my roommate being a technical representative (i.e., salesman) from one of the leading data processing hardware manufacturers in the country. During the course of a heated session (i.e., the room temperature was up and the speaker boring), my roommate uttered an appraisal of the program by stating that while he was not learning much new at least the speakers and participants were using the right set of "buzz" words in their presentations, discussions, and conversations. In my best academic manner so that any naivete and innocence would not be betrayed, I cautiously inquired as to what he meant. My companion replied that in almost any professional field there is always a current set of terms that one must use in his speeches, cocktail conversations, and business dealings to show that he is *with it* or *in*. It was comforting to me to learn that fields other than education have a problem similar to the one with which we are faced.

Education is beset with its own set of "buzz" words plus those from outside of education. The list of such words is relatively long and hence I should not use the time allocated to me to cover all of them. For those of you not familiar with some of the terms and/or the "buzz" word game, I would call your attention to an opportunity of being initiated which is presented in the September issue of the *Phi Delta Kappan*[2] wherein there is an opportunity to devise your own set of buzz words on a somewhat random basis.

I am sure that the symposium which you are attending has already presented, and will continue to present, its contribution to the pool of "buzz" words. The purpose of the symposium is not, however, to introduce new words into the language of the educational situation. Instead, it is a recognition of the fact that certain words have already become part of the

---

[1]A paper presented at the Symposium on Operations Analysis of Education, sponsored by the National Center for Education Statistics, U.S. Office of Education, November 20-22, 1967, Washington, D. C. Appreciation is expressed to Donald Miller, Roger Kaufman, Ed Novak, and Duane Dillman for their many helpful comments and suggestions.

[2]"Naming 1,000 Educational Innovations," *Phi Delta Kappan* (September, 1967), 36.

educational language. It is necessary, therefore, to give some attention to both the clarification of such words and to the possible role that the realities represented by the abstract word might have in the present and future of education.

In this paper, I propose to deal only with a selected set of "buzz" words derived from the more general one called Operations Analysis (working on the perhaps tenuous assumption that it is the *generic* term). The specific words that I would like to discuss are those of *project planning and control system analysis,* and *management techniques.* My single purpose is to indicate how these concepts can be related to and used in an activity which has been and is consuming more and more energies of professional educators—the activity is the preparation and execution of research and development projects.

DEFINING PROJECT PLANNING AND CONTROL

What is a *project?*[3] In view of a lack of consensus about a single definition, Gaddis provides a useful one for our purpose:

> A project is an organization unit dedicated to the attainment of a goal—generally the successful completion of a development product on time, within budget, and in conformance with predetermined performance specifications.[4]

While the above quotation provides the essential description of a project, the definition presented has a limitation for our use since it only relates to development projects. I should like to expand the definition to include the broader spectrum of research and engineering efforts as well as product development activities.

Projects of the above types have several common characteristics or features.[5] In general, projects can be said to be *finite,* or having a definite end point; *complex,* in that a mix of human and material resources is used to do a series of linear or parallel-related jobs; *homogeneous,* in the sense that one project can be marked off from another project or from the environment within which it exists (we will return to this characteristic later in this paper); and *nonrepetitive,* in that it is usually a one-time effort and therefore often has some uncertainty associated with it.

Planning is used here in its most general sense as described by Emery[6], that of outlining the future and/or deciding in advance what is to be done.

---

[3]Much of that which can be said here about projects applies also to programs. Major differences might be in scope, magnitude and duration.

[4]Paul O. Gaddis, "The Project Manager," *Harvard Business Review* (May, 1959), 89-97.

[5]IBM Application Description, Project Management System 360, Program Application H20-0210-0, 1966.

[6]James C. Emery, "The Planning Process and Its Formalization in Computer Models," *Second Congress on the Information System Science,* ed. by Joseph Spiegel and Donald Walker (Washington, D. C.: Spartan Books), 1965.

The output of the planning process is a plan which may take the form of a budget, rules, programs, schedules, and similar items. As Emery notes, the purpose of a plan is to bring about behavior that leads to desired outcomes. To accomplish the latter, the plan must (a) describe *actions* (some synonyms are procedures, process descriptions, and activity specifications) and *outcomes* (some synonyms are declaratives, state descriptions, or product specifications) and (b) serve as a formal tool or vehicle for management.

Control is used here in its most general sense of a monitoring function to make sure the plan is being effectively and efficiently carried out. The control formula of noting deviations from plan, taking necessary corrective actions, and recycling is included within the concept for purposes of our present discussion.

DEFINING SYSTEM ANALYSIS

The concept of system analysis presents some difficulty in being defined because of some confusion existing in the field at the present time. Let me start by noting, first of all, that the term *system* is being used in its singular and not in its plural sense. There is much confusion today in education, and perhaps even in other fields, relative to the use of the plural "systems" as an adjective and/or noun. Educators talk about systems approaches, instructional systems, and even systems analysis, but one is not sure whether the singular or plural is being intended.

System as used here refers primarily to the orderly (i.e., logical) arrangement of interdependent components or parts into a connected or interrelated whole to accomplish a specified goal. So defined, it is assumed that a system can be factored or resolved into a series of subsystems and each subsystem itself can be further factored or resolved.[7]

For a meaning of *analysis*, the ideas presented by Starr in his book on production management[8] are useful to us. Starr defines *analysis* in terms of the *principle of disassembly*. Under this principle, analytic behavior consists of operations that involve division, dissection, classification, partitioning, and similar actions. The operations of summation, integration, unification, and similar actions that relate to a principle of assembly, are the act of *synthesis*. It will be reassuring for many of my educational colleagues to know that synthesis has elements of Gestalt psychology associated with it.

Combining our concepts of *system, analysis,* and *synthesis* we can now define *system analysis* as that process of disassembling some objective-oriented whole into its component parts. *System synthesis* consists of putting the parts back together again into some kind of a whole. Figure 1 shows these two concepts in a graphic manner.

---

[7] David G. Ryans, "System Analysis in Education," *Technical Memo* (July), 1964.

[8] Martin K. Starr, *Production Management: Systems and Synthesis* (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964).

Figure 1

SYSTEM ANALYSIS AND SYNTHESIS

System Analysis	System Synthesis
Disassembly into:	Assembly into:
Parts	Wholes
Units	Entities
Subsystems	Networks
Activities	Total Systems
Functions	Flow Diagrams
Tasks	

DEFINITION OF MANAGEMENT TECHNIQUE

Management technique is not quite so difficult to define as the previous terms. Any attempt to define what is meant without a prior reference to the nature of management would be somewhat premature. Basically, management can be considered as a process which involves the functions of planning, organizing, directing, and controlling the personnel and other resources needed to accomplish an objective or goal. It is generally recognized that a manager's principal role is to make decisions with regard to each of the functions noted above.

Many techniques (or systems) have been developed in order to make the manager's task of carrying out the above role an easier one. In their recent book on management systems, Archibald and Villoria provide a useful definition of such a system:

> We may define a management system as a set of operating procedures which personnel carry out to acquire needed information from appropriate sources, process the data in accordance with a pre-programmed rationale, and present them to decision makers in a timely, meaningful form. Most contemporary systems involve manual data collection and input, machine processing, tabular and graphic output production, and human analysis and interpretation. Thus we can say that the systems collect, synthesize, process, transmit, and display information which flows from a primary source, through an editing, computation, and selection process to the manager.[9]

Two principal ideas are highlighted in this definition. First, management systems are designed to provide information. Second, the decision-making operation is left to the human manager. Some sophisticated systems do have what is called pre-programmed decision-making (i.e., rules) as an inherent part of them. Most systems, however, still rely upon the use of humans to make what in effect are non-programmed decisions. Most of the systems

[9] Russell D. Archibald and Richard L. Villoria, *Network-Based Management Systems* (New York: John Wiley and Sons, Inc., 1967).

developed are primarily aimed at facilitating the control function of a given manager's job. In view of the definition and ideas presented above, it is perhaps more appropriate to talk about management information systems rather than management systems. There is a general consensus that managers operating within the project context require data or information relative to time, or schedules; cost, or resources; and performance, reliability, or quality of objective accomplishment. Of these three types of basic data, the most common data obtained and used in a project situation is that relating to time or schedule.

A wide variety of numerous management systems or techniques have evolved over the past several decades to facilitate the manager's task. Three general types of systems have been developed. One group relates to the quality characteristics of a product, a second category relates to the operations involved in producing the product, while a third group relates to the administrations involved in carrying out the operations. The selection of a particular technique for a particular situation is not easy because a system designed for one purpose may not be suitable for a different purpose. I have chosen to delimit my definition of management techniques to that group of selected operations—related systems known as network techniques which have become increasingly popular during the past decade because of their relatively high degree of success in carrying out research and development activities with the military-industrial complex of our society. The application of such techniques to the field of education is just beginning, but their value has already been demonstrated.[10]

To summarize briefly, definitions have been provided for three current buzz words from the general area of Operations Analysis in education. *Project planning and control* was defined in terms of pointing out a concern with the planning or outlining of the future for a goal-directed activity which is finite, complex, homogeneous, and nonrepetitive in nature, plus monitoring and correcting deviations from the plan. *System analysis* was defined in terms of the analytic procedure involved in disassembling a goal-oriented whole into its component parts and then re-constituting it or synthesizing. *Management techniques* were defined in terms of their role of providing managers with information needed to control the accomplishment of the project effort within established time, cost, and performance parameters.

TOWARD PROJECT IMPROVEMENT

Having defined the three essential terms of concern in this paper, albeit briefly and simply, let me now present or develop the major point or thesis of this paper. It is contended that a sizeable number of projects in the field of education have been inadequately planned in their initial development and improperly executed once started because the initiators of such projects

---

[10] Desmond L. Cook, *PERT: Applications in Education*, Cooperative Research Monograph No. 17, OE-12024 (Washington, D. C.: U.S. Government Printing Office. FS5.212:12024, 1966).

were just not simply aware of the new tools and techniques available to them for project planning and control. It is further contended that the combined applications of system analysis and management techniques would be of immense value in producing better planned and controlled educational projects than has been the case in the past.

The basic premise for this position set forth is that the typical research, development, or engineering project in education can and should be fundamentally thought of as being a *system*. Viewing a project as a system can be justified primarily on that project characteristic referred to earlier as *homogeneity*. Any project can be visualized as being made up of a series of parts, units of activities which belong *exclusively* to that project. Hence, one can speak of the project as being an entity possessing edges and boundaries that help to distinguish it from the rest of the environment within which it operates. One writer put it, ". . . a project has a greater density of dependency *within* itself than *between* it and its surroundings."[11] The suggestion is made further that the boundary zone can be approximated where there are relatively few activities leading outward—that is, where there seems to be relatively natural perforation. Even though it can be distinguished from the rest of the environment, the project still operates within an environment which affects it or which the project itself affects. Under this definition or characterization, the concepts of system analysis and management techniques assume validity and become useful tools for project planning and control. The question now is, "When, where, and how are they applicable?" To answer this question, let us first examine the general steps in planning and controlling a project.

The initial step is to establish the goal or objective. The subsequent steps are to do a project definition, develop a project plan, and establish a schedule. The second step of project definition is essentially a process of disassembling or breaking out the many jobs which have to be accomplished to reach the stated objective. Putting it another way, we do an analysis of the system—or a system analysis. The end product of this definition or analysis step usually takes the form of a hierarchical plan or chart showing several levels of prime and supporting functions and tasks[12] (i.e., objectives) which have to be accomplished in order to accomplish the goal of the project. Functions or tasks can be further factored so that the smallest unit for planning purposes is established. Project definition through system analysis should not be thought of as an easy task. One difficulty is a clear establishment of that set of hierarchically ordered tasks or functions which helps to establish the boundary between the project and the environment within which it will operate. Another difficulty is the inability to define the measurable goals of the project. If the goal is a product, then the process to be defined by the analysis to achieve the product becomes quite tenuous if that original product is defined loosely.

---

[11]William H. Huggins, "Flow Graph Representation of Systems," *Operations Research and Systems Engineering* (Baltimore: John Hopkins University Press, 1960).

[12]The hierarchical plan could be established using variables other than tasks or functions, but the latter are the ones fairly commonly employed.

Having defined the project by use of analytic techniques, the project plan can now be developed by employing a graphic representation of the order in which the many functions or tasks have to be accomplished in order to reach the project objective. For this purpose, we can capitalize upon one of two methodologies system personnel have developed for representing any given system under consideration. One methodology involves the use of mathematical models or equations. While highly useful and very sophisticated to some system analysis, the set of equations has limitations in that it fails to portray in a readily comprehended form the structure of a system as a whole.[13] The second methodology for a system structure representation takes the form of various flow graphs, more commonly called "flow charts," "block diagrams" and similar terms. The purpose of such diagrams is not necessarily to portray the things that comprise the system but rather to show the various operations that the system performs upon the stuff it processes. Huggins has referred to flow graphs as a kind of mathematical Esperanto for system analysis, design, and simulation. They have the advantage of permitting a human to perceive quickly a total pattern plus the relationships existing therein and of being more palatable to the relatively unsophisticated person than are mathematical models.

The flow graph model which appears to have high relevancy for system representation is that one composed of a network of *branches* that connect at *nodes* or *points*. Each branch may be thought of as originating at one node and terminating at another node with direction from one node to the next point being indicated by an arrow. For our immediate purpose, each of these branches can be equated to or identified with the functions or tasks which must be accomplished to reach the goal.

Flow graph methodology permits one to achieve a synthesis of the functions identified through analysis. This technique helps to assemble that pattern of relationships, dependencies, and sequences which might be established as the most efficient way of "moving through the project." This "moving through the project" refers to the idea expressed earlier that flow graphs show how we intend to process the "stuff" that goes through the system. In a project, this "stuff" is most probably the intellectually-related activities required to accomplish the project objective. We are not able to portray graphically the intellectual processes involved in any of the functions or tasks, but we can show how we intend to organize and sequence our thinking and related activities as we move through the project in so far as the tasks and functions identified reflect such intellectual or cognitive processes.

Within system analysis techniques, a "functional flow diagram" is often employed to show the functions to be performed and the flow of these functions in meeting the objectives which have been identified.[14] The functional flow diagram has limited use as a management tool for making sure that the

---

[13]Huggins.

[14] R. A. Kaufman and R. E. Corrigan, *What Is The System Approach, and What's In It for Administrators* (Orange, California: Chapman College, 1967).

several functions and tasks are accomplished during the execution phase of the total project effort.

The concept of flow graphs to represent a system is somewhat of a recent origin. A very similar idea, however, was advanced by Gantt in the development of bar charts during the first World War.[15] Gantt charts served a useful purpose, but they had the limitation of not fully integrating the parts of the system into a component whole. Partial analysis was achieved through Gantt charts since the diagnostic synthesis perhaps served as a forcing function on what was analyzed. The analysis did, however, usually miss the interaction relationships. Other techniques using a visual or graphic representation of a function or task breakout were developed subsequently to the Gantt chart, but time does not permit a discussion of them. The most recent innovation has been the conceptualization and development of management systems which utilize the system analytic and synthesizing procedures combined with the flow graph concept of nodes and branches. The general terms for such systems are network techniques; PERT/CPM are the most commonly known systems.

Network systems such as PERT/CPM usually require, as a first step, that the project objective be identified. Then the elements of the total project are identified and placed in some type of hierarchical order. This essentially analytic procedure results in a product most generally known as the "workbreakdown" structure. Within network techniques, the workbreakdown structure most often represents the end products (either in the form of hardware or software) of tasks or functions rather than a representation of the tasks or functions themselves. There is no reason, however, to believe that such a product orientation must always obtain. Workbreakdown structures can be composed of tasks and functions just as easily as of products. In either case, the workbreakdown structure represents the project definition phase of total project planning. Based upon the workbreakdown structure, a network consisting of activities or functions (our original branches) and events (our original nodes) is developed with a uni-directional flow (usually from left to right). The network integrates the constituent elements of the project into their necessary sequence and dependency relationships (interactions) in much the same way as the functional flow diagram does in system analysis. Once the network is established, PERT/CPM have provided procedures for securing time estimates for the completion of each of the several activities or functions. These activity time estimates are then used to derive estimated start-completion times for the start or completion of functions, including the final node in the network. Schedules are then developed for the accomplishment of events or activities after careful consideration of resource availability and/or allocation.

Once the project definition phase involving analysis has been completed, the workbreakdown structure prepared, the process of synthesis accomplished through the network or flow graph procedure, and the time estimates/schedule procedure completed, the essential steps of project planning

[15] Wallace Clark, *The Gantt Chart* (New York: The Ronald Press, 1922).

have been accomplished. The result is a plan for the future—a graphical representation of the project tasks as they are to be accomplished in order to achieve the project objective. It should be noted here that we have approached the problem as if there were only one way to reach the project objective. Actually, there are alternative ways of reaching the goal. Each alternative with its associated project definition, project plan, and schedule is a separate system. If we prepare several alternatives for consideration, each may be considered as a separate system and thus we are approaching the *systems analysis* concept developed by Hitch and McKean.[16] Preparation of alternative systems to reach the project goal for management decision requires information relative to the associated costs for each system.

The network with its associated schedule provides us with a means of managing the project to make sure that the several functions or tasks are being completed in proper sequence and that we are staying on our time target. If the progress through the project is normal and on schedule, we can feel safe in assuming that the initial analysis and synthesis operations, plus the time estimating procedure, have been fairly well done. Should we begin, however, to observe malfunctions in the system, represented by nodes or events not being reached on schedule, or there is frequent modification of the plan in that new tasks are continuously inserted into the flow graph, such an observation might indicate that there was perhaps a failure to do a complete and thorough analysis and synthesis operation in the initial development of the project plan. Under this situation, valuable project time would be lost as resources were redeployed to conduct the necessary replanning operation. A schema illustrating the suggested relationships among system analysis, management, and project planning and control steps is presented in Figure 2. The middle box shows the steps in project planning and control while the relationships to system analysis are shown on the left and to management techniques on the right. The schema points up rather clearly that there is a great deal of similarity between the generalized sequence of steps under system analysis and synthesis and the steps of workbreakdown structuring and networking under network based management systems. Putting the results of the above steps into a time frame puts us into the realm of management control systems.

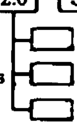
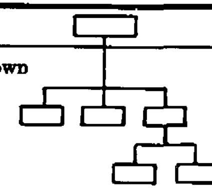
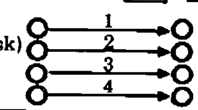
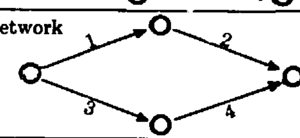
SUMMARY

The purpose of this paper has been to show how two current buzz words employed in educational circles—*system analysis* and *management techniques*—might be employed in combination to produce better planning and controlling of educational research and development projects. Experience with the use of these techniques has demonstrated that better

[16] C. J. Hitch and R. N. McKean, *The Economics of Defense in The Nuclear Age* (Cambridge, Massachusetts: Harvard University Press, 1960).

Figure 2

SCHEMA FOR INTEGRATING SYSTEM ANALYSIS AND MANAGEMENT SYSTEMS WITH PROJECT PLANNING AND CONTROL STEPS

System Analysis	Steps in Project Planning and Control	Management Systems (Network Based - PERT/CPM)
A. System Objective - Purpose	Project Objective(s)	Objective
B. System Analysis by hierarchical factoring into levels of functions Profile: [1.0] [2.0] [3.0] Functional Analysis 	Project Definition in terms of subsystems, products, functions, and so on.	Workbreakdown structure  Activity (task) definition (means-end) 
C. System Synthesis through flow graph methodology (functional flow diagram) [1.0]→[2.0]→[3.0]	Project Plan	Network 
(No Provisions)	Scheduling	Activity time estimates, event times, slack, resource allocation, etc.
D. System Feedback Loops Iteration	Controlling-Evaluating	Deviations, Management Actions, Decisions. Recycling

planned and controlled projects result than might otherwise be the case.[17] The employment of network techniques, such as PERT and CPM, for project planning and control, not only have advantages in and of themselves, but they also provide a very useful vehicle for an initial orientation into some rather complex operations—those of system analysis and synthesis procedures and the representation of existing and proposed systems through flow graph methodology.

The temptation to employ just one more buzz word cannot be resisted. The essential position in this paper is that the combined use of system analysis and management technique will have a *synergistic* effect upon project planning and control. As I understand synergistic, it is the effect of the two techniques operating in combination to produce an effect that either one would or could not produce by itself. The simple contention of this paper is that the synergistic effect from using system analysis and management techniques in project planning and control cannot be other than beneficial provided that such procedures are fully understood and properly applied. While not curing all ills, they can go a long way in helping with the stresses and strains exhibited under a bad case of "projectitis" under such programs as ESEA Titles I and III.

Let me close with a quotation which appears to be relevant both to the intended tone of this paper and the general topic of the symposium. It is a quotation from *Alice's Adventures in Wonderland* by Lewis Carroll.

> 'Speak English,' said the Eaglet, 'I do not know the meaning of half those long words, and what's more I don't believe you do either.' And the Eaglet bent down its head to hide a smile; some of the other birds tittered audibly.[18]

---

[17] For example, the joint work of the author with that of R. Kaufman, B. Corrigan, and D. Miller on the use of systems analysis and synthesis techniques and management systems in preparing educational planners for the state of California.

[18] Lewis Carroll, *Alice's Adventures in Wonderland* (New York: St. Martin's Press, 1965).

THE NATURE OF PROJECT MANAGEMENT[1]

Desmond L. Cook
Educational Program Management Center
The Ohio State University

INTRODUCTION

The role or position of project manager, also referred to as project director or principal investigator, is a relatively new role not only in the field of education but also in other areas, such as the military, government, and business. No specific point marks the introduction of this role. It has been gradually emerging in the last two decades as a consequence of increased mission-oriented activities which in large part have been supported by federal funds. As a consequence, there has been an emerging concern with the more general concept of program management. The purpose of this chapter is to discuss the general concept of program management, present characteristics of projects, discuss the general approaches to placement of projects in organizations, outline some of the problems faced by the project manager, and to indicate some of the factors or conditions which make for a productive project management.

What do we mean by the concept of project management? An answer to this question might be found by examining definitions which have been presented by persons who have done some writing in this field. Baumgartner[2] defines project management to "... consist of the actions involved in producing project deliverable items on time, within the contemplated cost, with the required reliability of performance ...." Cleland[3] defines project management as "... the means of managing a large aggregation of resources across functional and organizational lines of authority." Gaddis[4] discusses project management in terms of the project manager's ability to use the brain power of professionals and specialists in the creation of a product from its initial conceptualization through testing through production. The above definitions can be summarized to indicate that the project manager's principal role is to produce a product by integrating professional persons into a team operating within time, cost, and performance parameters with that team operating within some lines of organizational responsibilities and authority.

Such definitions, while useful to provide a general orientation, are not sufficient by themselves to fully explain the nature of project management

[1] Preliminary draft of a chapter for a book, *Educational Project Management*, to be published by Charles E. Merrill Co.. Reproduction is prohibited without permission of author.

[2] John S. Baumgartner, *Project Management* (Homewood, Illinois: Richard D. Irwin, Inc., 1963).

[3] David I. Cleland, "Why Project Management?" *Business Horizons*, VII (1964), 81-88.

[4] Paul O. Gaddis, "The Project Manager," *Harvard Business Review*, XXVII (June, 1959), 189-97.

or the role of the project manager. We need, therefore, to take a look at the concept of a project as well as the nature of management, the organizational placement of projects, and related topics. The emphasis in the present chapter will be primarily about the nature of project management and its role in the organization. Subsequent chapters will deal with the general nature of management.

CHARACTERISTICS OF PROJECTS

An understanding of project management requires some orientation as to the general nature of projects, and thus characteristics, which help to distinguish such activities from what might be called non-project activities. Gaddis[5] defines a project as ". . . (an) organizational unit dedicated to the attainment of a goal—generally the successful completion of a development product on time within the budget, and in conformance with performance specifications." Woodgate[6] defines projects as ". . . work to be done or procedures to be followed in order to accomplish a particular (project) objective." Robertson[7] defines a project as ". . . the accomplishment of a number of actions in series and/or parallel in order to reach an objective." An examination of these definitions can lead us to establish some common elements existing among them as a means of being able to characterize project activities from non-project activities. Four characteristics are presented below which can help us in this process.

First, projects are usually finite in character to the extent that there is usually only a single objective to be accomplished and the project terminates upon the accomplishment of the objective. Further, the objective to be accomplished is usually set within some time, cost, and performance specifications. This characteristic should not be interpreted to mean that only a single objective is involved, because in many cases multiple objectives might be involved. The main concern here is that at some point in time the objective or objectives will be accomplished, at which point the project is terminated.

Second, projects are usually complex in nature to the extent that they involve a large number of tasks to be accomplished by personnel and other resources to accomplish the established objective. Projects can differ in the degree of complexity involved. Projects of relatively short duration, say, three months in length, designed to accomplish a very limited objective would not be so complex as a project lasting perhaps three or four years and culminating in the achievement of several objectives with a large application of resources of various types. No specific criterion has been established with regard to the degree of complexity. The principal point here is that there is a relatively large number of individual tasks which have to be done by people in order to accomplish the goal.

---

[5] *Ibid.*

[6] H. S. Woodgate, *Planning by Network* (London: Business Publications Limited, 1964).

[7] D. C. Robertson, *Project Planning and Control* (Cleveland, Ohio: Chemical Rubber Co., 1967).

Third, a project consists of a series of tasks which relate only to that effort and, therefore, it is possible to distinguish the project from the environment in which it exists and also to distinguish it from other projects which might exist in the same environment. Each project consists of a unique set of tasks which relate only to that effort. It is possible, therefore, to conceive a boundary line between the project and the rest of its environment. We can say that a project is more alike within itself than between other projects in the environment around it. One of the major problems is the establishment of boundary lines so that only those tasks that are related to the effort can be identified. It is helpful in this regard to think of the project as a system and, therefore, amenable to the concepts, principles, and procedures of systems analysis. We will return to this consideration in Part II.

Fourth, a project generally consists of a once-through, non-repetitive or a one-of-a-kind activity. A particular project will be done only once. Any one project will not be repeated even though other projects similar to it may be undertaken. For example, the construction of a particular school building can be considered as a once-through project even though other school buildings will be built. A particular school under consideration, however, will not be repeated. Research and development projects, by their vary nature, are once-through efforts. One of the major problems in such non-repetitive activities is that there is very low historical and/or objective information which one can go upon to establish the time, cost, and performance specifications as well as the specific tasks to be done. This concept of uncertainty pervades many small and large-scale research and development projects, is one of the major characteristics of projects, and presents a formidable management problem. Consequently, there is a need either to develop and/or utilize techniques in the management of educational development projects which help the project manager to deal with the uncertainty problem.

The determination of whether or not we are dealing with a project or non-project situation can be made by re-casting the characteristics presented above in terms of a series of questions. If we answer the questions in a positive manner, then our situation is one to which the techniques presented in this book are devoted. Should we not be able to answer the above questions in a positive manner, the chances are that the techniques presented herein would be inappropriate.

THE PLACEMENT OF PROJECTS IN AN ORGANIZATIONAL CONTEXT

A proper understanding of the nature of project management and the function and duties of the project manager require an orientation of the place the projects have in an organizational structure. In discussing this topic, an assumption is made that most projects will be housed or placed in an existing educational agency as opposed to being placed in a specially created independent agency. The latter situation might be true in the case of exceedingly large projects but such cases are not typical. A most typical situation is to have the project placed within an existing organization. Five general types of project placement will be presented below, and the relative advantages and limitations of each presented.

Separate Organization Approach

One approach is to place the project within the organizational structure but completely independent of any existing functional units or departments. In this case, the project is self-sufficient with regard to the project requirements. All personnel fall under the project manager's authority. Staffing is provided either by transferring personnel from other departments or by hiring from outside sources. Such a unit is normally dissolved upon the completion of the project, with staff members reverting back to their original departments and/or leaving the organization.

This approach has advantages in that the project manager has direct control over all of the dimensions of the project and the main line of authority is principally to the chief executive of the organization. One of the major limitations in this approach is that the project staff assembled for conducting the project will be lost upon completion of the project. In such instances, the organization, if not able to place persons in other areas, can lose the services of some very valuably trained specialists.

Vertical or Centralized Approach

A second general approach to project placement in the organization is to locate it in a vertical line unit or department. All the necessary project staff are drawn from that one department and placed under its administrative head. In this kind of an approach, the lines of authority and responsibility are known to all persons working on the project. The functional department may or may not have the staff to work on the project, in which case individuals may be drawn from other functional units and placed in the department for the duration of the project. Persons may be employed from the outside to work on the project. This approach has the advantage of well-known lines of authority and responsibility. It has the limitation that such arrangements can lead to new and powerful units as individual projects increase in size and number. This approach often results in "empire building." If the empire becomes big enough, the organization may find itself moving a sub-unit from one department to a level of a major department. This approach appears not to be a common one employed in many agencies because of the problems of securing competent staff from within the department and of "empire building."

Horizontal or Decentralized Approach

The most typical approach with regard to project placement is to superimpose it upon an existing vertical and horizontal structure with known lines of responsibility and authority. In a sense, the project becomes a horizontal unit working across the vertical organizational structure. Personnel needed to work on the project remain within their department or functional units. Project tasks are assigned to functional departments as required. The project manager is responsible for the completion of the work and generally has authority over the *what* and *why* of the project, while the functional department manager usually retains authority over the *how*. The major limitation to this approach is that the project manager often lacks authority over the vertical functional areas in which the project tasks are accomplished. Because of this situation, the large single purpose project which cuts across functional organizational lines of authority causes unique management relationships and requires a new and different management philosophy than has existed in the past. The traditional functions of superior-subordinate lines of authority are difficult to apply in this approach when the project involves a coordination of a large number of persons working under departmental heads whose authority and responsibility lines are vertical in nature.

The Executive Staff Approach

In some cases, one or more projects may exist in an organization, perhaps in functional units, and one person is designated to serve as a staff assistant to the chief executive. This person's responsibilities are to coordinate, analyze, and make recommendations regarding project situation, but the final decisions are made by the chief executive. Under this arrangement, the person placed in direct charge of the project cannot function effectively as an integrator and decision-maker since these have been taken out of his hands. This type of approach is very similar to that of the role of the "expediter" in many organizational units.

Project Staff Approach

This arrangement usually consists of a project manager being appointed with a staff available to carry out such processes as scheduling, task and cost control, and other special functions unique to the project. The functional departments still perform the majority of the work associated with the actual completion of the project. Such an approach requires considerable coordination between staff and functional department. This approach is very closely allied to the decentralized approach.

As was noted, a most typical approach is to place the project within an existing organization with the assignment of a person or staff to be in charge of the project. In such situations, certain general problems present themselves which make the job of project management quite difficult. It often happens that when functional departments are more interested in their own specialties than in contributing to a unified project effort, oftentimes, the total perspective of the project is lost. Actions are taken without regard to their effect on the total project activity. Often, decisions pertaining to the project are slow and difficult to make and required information relative to problems must be obtained through several channels from the various functional departments. The very nature of projects often requires rapid adjustment to new situations. In many cases, functional departments are not able to be sufficiently flexible to deal with such adjustment. It is these kinds of situations and problems which have created an interest in the topic of project management in order to develop solutions to them.

RESPONSIBILITIES OF THE PROJECT MANAGER

With perhaps the exception of the executive staff approach, the responsibilities of the project manager within the other approaches generally tend to be the same. Some of these responsibilities are outlined below.

The project manager is the focal point for the total project effort. Most of the major project considerations relative to time, cost, and performance are channelled through his office. He is responsible for providing a means for integrating and systematizing decisions, policies, and managerial priorities for the various functional and organizational elements. He thus becomes personally involved in critical project decisions. Since he is working with professionals in functional departments as well as with higher levels of management, he serves in a role as a "go-between" for these two groups. Consequently, he must be at home not only talking about the technical research and development problems on the one hand, but also being able to talk about management problems and concerns on the other.

The project manager is responsible for developing the project plan. He must recognize that advanced planning is the vital part of effective project management. He must be able to identify fully the objectives of the projects as well as the major events or completion points. The amount of detail and realism of the project plan are also one of his major concerns. The project manager must be aware of the fact that most of the crises or problems that develop in a project usually occur because of a lack of or the absence of advanced planning.

The project manager must keep control of the project at all times. Schedules, budgets, and performance specifications must be put together in a meaningful way so that he knows what actions to take and when to take them. He must recognize that in his role he is probably the only person in the project who can integrate the *what* and *why* with the *when* and *how* of a particular task. This problem of control becomes very acute because he must be concerned with the management and conservation of funds entrusted to him under contractual arrangements with the funding agency.

The project manager has the responsibility for developing the project staff. Necessary personnel with the requisite skills must be secured at the proper time and place. The team spirit existing or not existing among the members of the project staff may be a consequence of a lack of concern with this responsibility. It will be upon his initiative that the managerial dimensions of the project team will be developed as needed in order to facilitate successful completion of the project.

The several possibilities outlined above cannot be delegated to other persons on the project. Certain operational activities might possibly be delegated, such as the recruitment of personnel, but the final decision regarding such activities is the responsibility of the project manager or director.

PROBLEM AREAS IN PROJECT MANAGEMENT

Studies of the emerging role of project managers have identified certain problems which exist with which a prospective project manager should be concerned. Some of these problems are presented in this section.

Because most research and development projects involve the work of a number of professional persons with specialized skills, the traditional superior-subordinate role becomes inappropriate in working with such persons. Such professionals are often more concerned with how a task is accomplished and are in often conflict with the project manager who is more concerned with the what and when of task accomplishment. He must recognize that professionals are often being asked to produce on a schedule and be aware of the implications of this for their professional modes and manners of operation. He must also recognize the fact that many professionals like to work toward perfection but that in doing so the total project effort might be delayed.

As noted above, most projects are finite in a time dimension. Hence, the assignment of the project manager will terminate unless the project is renewed. Impending project termination may cause a project manager to become more concerned with his future status toward the end of the project and thus begin to spend more time preparing new proposals than in managing the present situation. As a consequence, final activities associated with the project may be neglected as the project manager devotes his time and energy to the proposed new effort. Because of this situation, educational institutions and agencies must begin to develop some procedures or systems which will give reasonable assurance to the project manager that some position continuation will be available upon project completion or until subsequent funding is obtained upon completion of the immediate project.

One of the major problem areas was noted above when it was pointed out that the typical project manager has few if any lines of authority over staff members in other departments. As a consequence, other techniques or manifestations of authority have to be employed in order to accomplish the project tasks. The project manager must draw upon his persuasive abilities, reputation, rapport, influence, status, and prestige in order to accomplish the project tasks. He must also be able to make use of lateral agreements made between the project and other organizational units. The basic authority that the project manager does have lies within a type of *de facto* authority grant with the awarding of project funds. His ability to use these funds as a means of authority becomes an important part of project management.

In some cases, projects are placed in organizations in a manner such that the project director does not have effective control over the schedules, budgets, and performance dimensions of the projects. Decisions regarding these three dimensions may be placed in the hands of other functional unit department heads. It is generally agreed by specialists in project management that this is an undesirable arrangement. The inability of the project manager to develop any meaningful relationships among these three dimensions of a project, because decisions regarding them are made by other persons, may mean that corrective actions may not be able to be taken when appropriate. Placement of these three dimensions under the direct control of the project manager gives assurance that there are plans, controls, and an organization for getting the job done. This may mean that the project director be given a necessary executive rank within the organization to ensure that the parent organization will pay attention to this request. It may be helpful in this regard to provide orientation meetings to existing organizational structures and units regarding the role of the project manager in order that he can secure their cooperation.

One of the major problem areas centers around the undue concern on the part of the project manager with trivial details of the project. His primary concern should be on operating on an "exception" basis. His time and energy should be devoted to the identification and solution of critical problems of the project and not with the minute details or non-crucial problems. He should be aware of the fact that the absence of problems may be just as indicative of trouble in the project as continuous crisis situations.

While focusing upon the problems facing the project manager described above, we should recognize that the long range developments of the organization and personnel may be hindered if there is only concern with the management of projects. Priorities assigned to projects may upset the stability of the organization and interfere with the accomplishment of its functional task. Further, the shifting of personnel from project to project may disrupt the training of new employees and specialists, thus hindering their growth. It is also quite possible that lessons learned in one project may not be communicated to another project. Project managers, therefore, must be always aware of the impact of the project on the organization and work not only to promote the efforts for which they are responsible but also to integrate their efforts with the total organization so that undue conflicts do not arise and the total organization moves forward.

FACTORS LEADING TO EFFECTIVE PROJECT MANAGEMENT

The successful project may depend more often upon the effectiveness of the project manager than upon the technical dimensions of the project. Some of the factors or conditions which have been identified as leading to effective project management are presented below.

The person assigned to the role of project manager should have an experience and background in those areas or fields of education which relate to the major substance or thrust of the project. No one person can be competent in all areas of education. He should have the fundamental knowledge which can be augmented when necessary in order to deal with the details of a specific operation. His educational background should not only relate directly to the field of education but also he must have a good understanding of management concepts, particularly such topics as planning and controlling decision-making.

In addition to an understanding of management topics as indicated above, the prospective project manager must be familiar with the general technology of scientific analysis and integration. He should be able to analyze a problem into its component parts and then integrate it back into the total effort. Training in the scientific method, therefore, becomes a very important tool in the preparation of project managers.

If it is at all possible to do so, any person anticipating assuming the role of project manager should have had some experience or time as a project manager or director in some previous project. Lacking direct experience as a project manager, participation in a project in the role of an associate project director often provides the needed background in order to assume the role of project manager. The experience gained as a research associate by participation in staff meetings where management problems are discussed often creates an awareness on the part of the person filling such a role to acquire training in this area.

The role of the project manager becomes so important that it would not be out of place to establish training programs which are designed specifically to develop the skills needed in the area of project management. Several efforts are already underway in this direction. It would seem desirable that training specifically in project management should be a desirable part of the training of persons who are interested in carrying out research and development projects in the field of education under the auspices of a funding agency such as the U.S. Office of Education. Such knowledge can be obtained by participation in short courses put on by various management associations and agencies. Educational institutions and agencies could develop short-term programs which could focus upon the problems and procedures associated with project management. The type of training indicated here would be devoted specifically to the role of the project manager. Participation in training courses dealing with the general nature of management is also valuable.

A last point to be made with regard to effective project management is that organizational support and resources must be provided in order to work successfully. The lack of top-level support can often be a disrupting factor in the success that the project manager has with his project. If he is not provided with time and assistance to carry out his many duties associated with the project, or finds himself involved in conflicting lines of authority, the whole effort can be jeopardized. It will appear only natural that, if the project manager is given the responsibility for the total effort, the organization within which he is working gives him the support needed to ensure the success of the project.

A GENERALIZED PROJECT MANAGEMENT SYSTEM MODEL

Desmond L. Cook
Educational Program Management Center
The Ohio State University

INTRODUCTION

The purpose of this paper is to present the outline of a generalized model for the management of educational projects. The model presented has been derived from a study of existing management systems, their functions and organization, experience in actual management of projects, conducting instructional programs on management systems, and an analysis of the existing literature on project management procedures and problems.

The generalized model presented does not represent any particular existant management systems, although many of the concepts of network-based systems have been incorporated.

It is anticipated that in presenting the generalized model an individual project manager would be able to develop his own project management system by utilizing all or part of the model, or modification of parts, as desired or needed for the unique situation. It is also hoped that, by presenting the model using selected systems concepts and principles, the reader can begin to become familiar with such concepts and how they are utilized in an actual application.

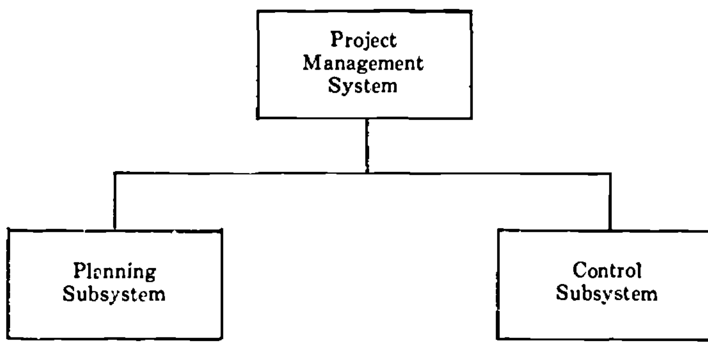
Comments and suggestions regarding the model are solicited.

MODEL PROJECT MANAGEMENT SYSTEM

Total System Function

The function of the project management system is to develop a plan for use by a project which includes time, cost, and performance specifications and to provide a vehicle for monitoring and controlling project plan operation once the project is initiated until completion or termination.

Accomplishment of the total system function is made by the identification and development of two major subsystems which relate to the two management functions of *planning* and *controlling*. Each of these two major subsystems is described in the sections which follow. The figure below outlines the two major subsystems. No attempt is made at this time to present or develop subsystems which relate to the management functions of *organizing* and *directing*. These subsystems will be added to the model at some subsequent time.

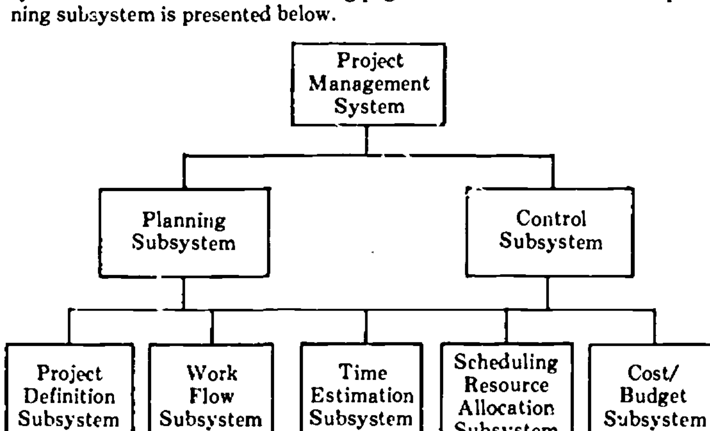


THE PLANNING SUBSYSTEM

Subsystem Function

The function of the planning subsystem is to provide (a) a plan including schedule and budget for accomplishment of the prime project mission or objective along with a supporting objective coupled with (b) a data or information base which can be utilized in the control function to identify problem areas (i.e., deviations from plan).

The accomplishment of the subsystem function requires the identification and development of functions and subsystems relating to (a) project definitions, (b) work flow, (c) time estimation, (d) scheduling and resource allocation, and (e) cost/budget estimates. Each of these functions and subsystems is discussed on the following pages. An outline of the over-all planning subsystem is presented below.



Project Definition Subsystem

Subsystem Function. The function of the subsystem is to develop the boundaries of the project by establishing a hierarchical structure of the major and subordinate objectives reflecting work that has to be accomplished to reach the over-all goal of the project and which are expressed as products or functions along with their performance specifications or criteria of accomplishment.

Input	Sequence	Output
1. Major project objectives 2. System concepts 3. Product/function decision 4. Limits and constraints 5. Performance specifications 6. Processes involved in establishing objectives	1. Establish mission statement or over-all project goal 2. Identify major end item to accomplish step 1 3. Develop subordinate tasks to major end items in step 2 4. Assign responsibility for work package development 5. Determine criteria for work package accomplishment 6. Review and revise 7. Reproduce	Project definition in form of workbreakdown structure in graphic or tabular form which is product or function oriented

Work Flow or Plan Subsystem

Subsystem Function. The function of the work plan subsystem is to develop a graphical representation of the sequence of the tasks and events necessary to accomplish the objectives identified in the project definition subsystem taking into account necessary interrelationships and dependencies.

Input	Sequence	Output
1. Project definition 2. Rules for work flow plan to be used (Gantt, Network, etc.) 3. Computer use decision 4. Milestone events 5. Task/event numbering decision 6. Event coding system (milestone, interface, etc.)	1. Develop over-all work flow using milestone events 2. Establish interface events 3. Assign responsibility for work package flow development 4. Develop detailed work flow 5. Check for loops 6. Check final logic of work plan. 7. Revise as needed 8. Adopt task identification scheme (event numbering) 9. Reproduce work plan	Work flow in chart form (network) in master and/or detailed form

Time Estimation Subsystem

Subsystem Function. The function of the subsystem is to provide managerial information regarding total project completion time plus time for the earliest and latest time for the initiation and completion of individual work tasks.

Input	Sequence	Output
1. Decision on deterministic or probabilistic estimating procedure 2. Network or work plan 3. Directed or schedule dates if known 4. Persons to provide estimates 5. Work sheets 6. Knowledge of estimating procedures 7. Knowledge of permissible replanning procedures	1. Assign final date if known to terminal task on master plan 2. Assign scheduled dates to milestones if known 3. Secure and record estimates for individual activities 4. Calculate expected activity times 5. Calculate earliest task completion time 6. Calculate latest allowable completion time 7. Establish critical path and slack 8. Adjust to directed and scheduled dates as needed 9. Finalize	Work plan showing initial terminal time, milestone time, critical path, slack paths and individual task times

Scheduling/Resource Allocation Subsystem

Subsystem Function. The function of the subsystem is to schedule plans for the project by translating planned schedule derived from time estimation subsystem into specific calendar dates for the initiation and completion of work compatible with resource availability and other known or stated constraints.

Input	Sequence	Output
1. Resource survey 2. Decision on schedule criterion 3. Calendar considerations 4. Organizational personnel policies 5. Contractor requirements 6. Final task and milestone dates, if known 7. Determination of time or resource constraint situation	1. Develop functional charts (Gantt or Bar charts) showing critical path, slack paths, and job sequence 2. Determine resource needs for task accomplishment 3. Adjust task flow as needed to fit resource availability 4. Check against criterion 5. Revise as needed 6. Translate to calendar dates 7. Issue work authorizations	Work plan adjusted to meet resource availability showing start and completion dates for task accomplishment consistent with directed completion date

Cost/Budget Estimation System

Subsystem Function. The function of the system is to generate a budget or future expenditure plan which provides for the necessary costs or resources needed to accomplish the project as outlined and established in prior subsystems and to provide a basis for future decisions as well as control of current expenditures.

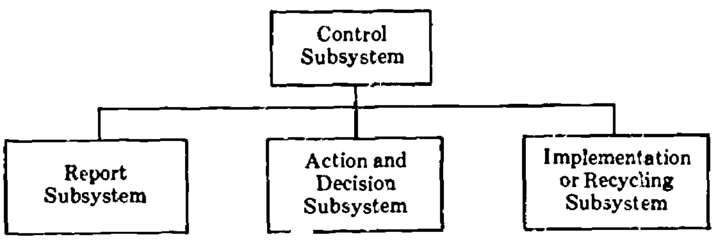
Input	Sequence	Output
1. Scheduled work plan 2. Project definition 3. Individual work tasks of work packages 4. Policies on costs (travel, etc.) 5. Knowledge of budgeting concepts 6. Negotiated indirect cost rates	1. Estimate costs for individual tasks 2. Estimate costs for work packages using information from step 1 3. Summarize costs upwards through project definition 4. Develop traditional category budget using data from step 1 and/or 2 5. Calculate direct and indirect costs figures 6. Establish planned rate or expenditure curves for work packages and total project	Budget for project showing costs and planned expenditure curves for major work tasks plus traditional budget categories

THE CONTROL SUBSYSTEM

Control Subsystem Function

The function of the Control System is to provide management with timely, relevant, and valid information so that problems (i.e., deviations from plan) can be identified, alternative corrective solutions considered and decisions made, and decisions implemented by recycling the project as needed.

To accomplish this over-all function, three separate functions and appropriate subsystems can be identified. These functions and subsystems relate to (a) report preparation, (b) management actions and decisions, and (c) implementation of management decisions or recycling. These subsystems are shown in the figure below.



Report Subsystem

Subsystem Function. The function of the system is to provide to appropriate management levels continuous, accurate and rapid detailed and/or summary information which reflects current project status and highlights present and potential problems in a form that is concise and clear.

Input	Sequence	Output
1. Data base from planning system 2. Report formats 3. Briefing techniques 4. Decision on management levels 5. Report dates 6. Updating procedures for organization 7. Deviation limits 8. Problem analysis skills and concepts 9. Duties and functions of network analysis	1. Note activities/events completed 2. Re-estimate as needed time/cost/performance for activities in progress 3. Re-evaluate time/cost/performance for activities not yet started 4. Process data 5. Analysis of outputs for deviations 6. Identify problems, show impact, develop alternatives with concern unit 7. Prepare visual reports with accompanying narrative 8. Forward to appropriate management levels	Periodic reports to management levels showing project status, problem areas, impact of problems, and suggested alternatives

Action and Decision System

Subsystem Function. The function of the system is to enable managers at various levels to develop actions and make those decisions which will resolve problems to correct deviations from original plans and/or to modify original plans as desired.

Input	Sequence	Output
1. Management reports 2. Knowledge of decision processes 3. Authority and responsibility for decisions 4. Contractor requirements if appropriate	1. Set problems priority 2. Identify and test causes 3. Develop objectives for decisions 4. Establish "musts" and "wants" 5. Develop and secure information for each alternative 6. Eliminate alternatives on go/no go basis 7. Study desirability of possible alternatives via simulation 8. Study adverse consequences 9. Make decision	Decisions reflecting solution to be employed to correct deviations or modify as needed

Implementation System

Subsystem Function. The function of the system is to provide a means of implementing management decisions, revising plans, and developing a modified data/information base.

Input	Sequence	Output
1. Management decisions 2. Methods of disseminating or communicating management decisions 3. Was/is worksheets	1. Transmit management decisions to appropriate unit 2. Modify original plans 3. Complete "was/is" worksheet 4. Process new data 5. Revise work plan as necessary on periodic basis	Revised plan reflecting adjusted time/cost/performance dimensions of present data

4
STATE EVALUATION, A TEAM EFFORT

by
Roy Jones

STATE EVALUATION, A TEAM EFFORT

Roy Jones, *Coordinator*
Division of Instruction
Tennessee State Department of Education

What is the place of program and/or project evaluation in education? What priority should it be given? Is it the concern of the research specialist only? We know that teachers are concerned, especially about certain aspects of evaluation. When faculty meetings become dull, just introduce the topic of *grading* and *promotion*. This will place the meeting in orbit every time. The superintendent who is "up for election" is interested in assessing the strength of his opponent. These are isolated examples; many could be cited.

Evaluation is not new to us. But I am afraid that well-structured, systematic program evaluation is new to many of us. In the past we have often relied on limited empirical data as the basis for decision-making.

Education in Tennessee is structured so that superintendents, board members, county courts, city boards, legislatures, congress, supervisors, principals, teachers, guidance counselors, students, and the community all make decisions that influence the quality of education. However, if I am reading the signs right in our state, legislators and the public are becoming extremely interested in how we invest the educational dollar. If you ask anyone whether he is for higher taxes, he will likely say, "No!" very emphatically. But if you ask whether he would be willing to pay more taxes for kindergartens, guidance counselors, and so forth, his reaction may be different.

With the legislature and public taking this kind of interest and with the authority to be selective in the specific programs they authorize, they can in effect take gross program planning out of the hands of the educator.

It is more and more imperative that educators be able to defend objectively the priorities that we are convinced will do the most at a given point in time to improve educational opportunities for the learner. This will require more than "seat-of-the-pants" judgment. It requires a well-structured, systematic evaluation program.

But, beyond this gross kind of decision-making, I am convinced that we must develop a learning environment and engineer learning activities that will make the learning process palatable and efficient for the learner. Perhaps we can say with some pride that we are nearer there than we used to be. But we would all admit that the step we have taken in this direction is much too small. We need the moon walk's "giant leap"—an educational giant leap for mankind.

We have the "know-how" to develop a viable curriculum utilizing organizational arrangements, methods, techniques, and content that give more opportunity to meet individual learning needs. We have many fine instructional materials and supplementary learning resources. We have the ability to design and build buildings that facilitate, rather than inhibit, the development of programs designed to meet individual differences and to promote independent learning. We have enough understanding of human relations to continually improve the social and learning climate of the classroom.

But to plan for these achievements educators must have relevant, reliable information which can be secured only through good evaluation techniques. Probably, evaluation is needed more than innovation at this point in time. Building on and refining that which is sound, throwing away that which inhibits learning, and innovating or inventing to fill in the gaps should result if evaluation is sound.

We may have to take a two-pronged approach to educational improvement, one of them based on the answer to the question: How do we make rapid improvements (short range goals)? and the other: How do we build or engineer a climate from which educational programs of extreme quality will emerge? Either of these approaches will demand the best information about our educational endeavors that we can secure. We will need information at the local system level, at the State level, at the university level, and at the Federal agency level.

Probably, we will be amazed at the amount of information we need. As we accelerate program building we multiply decision-making. This demands relevant, reliable information. We have a considerable amount of information now. But, if we can turn our evaluation efforts into team efforts with the local school system, State Department, universities, and Federal agency each contributing and receiving information, we believe we can get the most from our energies. Many needs exist in the area of evaluation:

1. A language that is precise and understood across the State.
2. A design that can be used successfully with any project or program improvement activity.
3. A structure for organizing information so that quantities of it can be easily studied.
4. More expertise in the area of data collection and treatment.

We are fortunate in Tennessee. We have a Testing Center at the University of Tennessee in Knoxville. Many of the local school systems use the services of this center to secure and score standardized tests. Through this procedure we have available a considerable amount of data. We have not utilized this data too well in the past, but we hope to in the future.

The source of data collection will of necessity be the local school. The responsibility for collection will fall on local school personnel. The State Department has assumed the responsibility of providing a design. We hope to provide as much assistance as possible because skilled evaluators are scarce. We do not want to develop a cookbook evaluation program, but we would like to provide a road map to help you see your alternatives and be able to capitalize on evaluation aids that have been developed by experts.

We have two evaluation teams staffed by Department personnel. Their background and training in the area of evaluation varies, but they are committed to helping local educational agencies in planning evaluation activities. They have been involved in the study of the design and will continue that study. They will be involved in State Program Evaluation of Title III, ESEA. As finances permit we hope to add more staff with training and experience in evaluation.

We hope that in the not too distant future our educational assessment file will have sufficient data to provide feedback to you which will be helpful as you plan projects or programs.

During the past several months, Memphis State University has made a great contribution to evaluation efforts in Tennessee. They will continue to assist us for the next few months as we refine the Department's role in evaluation. They will also assist in our first Title III State Program Evaluation, as well as in additional assessment activities.

We believe that the team approach is the best one that we can take with the local school system, universities, and State Department accepting roles that are reasonable. We think this approach will be profitable for the learner in our State.

We believe that making staff teams available to you for planning purposes and accepting a leadership role in the in-service training of local personnel in the area of evaluation is our best approach at this point in time. We believe a design that builds in evaluation planning as an integral part of project or program development is sound and will be one additional vehicle for the improvement of educational programs.

We still do not have all the answers in program development, nor do we have them in evaluation, but, as a friend of mine likes to say, "Our doin' hasn't caught up with our knowin'."

We must take a step at a time, but we hope we can lengthen these steps as we go along. If the climate and "know-how" permits, we should take a giant leap whenever the opportunity presents itself.

5
WRITING EDUCATIONAL OBJECTIVES

by
Donald R. Thomsen

WRITING EDUCATIONAL OBJECTIVES

Donald R. Thomsen, *Assistant Director*
Central Midwestern Regional Educational Laboratory
Memphis Area Office

The statement of an objective is a step in a strategy leading toward some identified goal which has been defined as a desirable, worthwhile purpose for American education. The strategy works something like this:

> The pupil must be able to identify integers so that he can name the integers so that he can add two integers so that he can multiply one number by another so that he can solve problems involving numbers so that he can meet the qualifications for a job so that he can pay taxes so that he can . . . .

The ultimate here may be the imperative of American education expressed by AASA: "To keep democracy working."

The other alternative for planning a strategy is to begin with the ultimate imperative and work backward to the immediate objective of the teacher. This merely changes the approach by saying:

> He can vote if he has a job. He will have a job if he qualified for the job. He qualifies for a job if he solves problems involving numbers. etc.

Either approach is appropriate, but both approaches are dependent upon specifying objectives. Each objective is stated in terms of what the pupil is doing when the objective is reached. How would it be possible to strategize without knowing what the pupil is able to do at each position in the strategy? Can you imagine paying a TV repairman to come into your home without your knowing he is able to repair your TV? You complain loudly if he comes in and spends hours—for which he expects to be paid—cleaning the set, rearranging the furniture, conversing intelligently, but leaves without the TV functioning properly. You expect him to perform a repair job only and your objective has been met when he leaves.

The pupil arrives at school. Should not the parents and taxpayers expect the same consideration for the teacher's time? Should they not know what the child will be able to do when the teacher has accomplished his mission? You are paid as a teacher to attain an objective for each child presented. It appears logical to me that you should be able to state what your objective is and whether or not the child has attained that objective. If you select the wrong objective, it is up to the citizens of the community, through the duly constituted board, to alter the policy of the school system that would lead to other purposes of education. The teacher is the professional strategist and, therefore, must establish the immediate and enabling objectives designed to meet a specified goal for education.

STATING EDUCATIONAL OBJECTIVES

How often have you heard a speaker begin by saying, "Webster defines this as . . ."? Let this not be an exception. It so frequently happens that we stereotype our speech with trite phrases or use words we hear others use without bothering to look at the root of the word itself. I have been amazed at how "modern" a dictionary is when I have looked up a word which has been used frequently and have discovered the new use of the word was actually a definition Webster probably used in his first edition. The word germane to this discussion is "objective" when used as a noun. Mr. Webster includes this: "the strategic position to be obtained."

Whenever you see an objective that declares the purpose of instruction is to "understand the use of the English language," as an example, have we stated that this is a "strategic position to be obtained"? Or have we merely defined a direction toward some goals, the description of which we do not clearly understand ourselves? By contrast, we may look at the word "goal." Back to Webster again—this is a terminal point or an intention. The words are almost synonomous, but the objective is stated as a position and the goal is a point.

Digging a little deeper, we look at the process of education as a process of changing behavior in a pupil. The change, therefore, is a describable difference in how a pupil acts or reacts. The goal for teaching English is to reach a point at which we are secure in the fact that a pupil will use the English language as a means of communication in a manner that we accept as proper. The objectives for teaching English may be the attainment of a position to which the pupil is prepared to move forward to the point defined as our goal. Thus, the objective may be in terms of using the proper sentence structure, the proper parts of speech, or familiarity with literary forms. This discussion may seem pointless and as an exercise in semantics, but I have a goal in mind based upon successive objectives which I believe you will attain.

* * *

Title III projects are funded upon the basis of a proposal you submit in which you are testing theory or instruction that is predictive of changes you will bring about in the behaviors of pupils or teachers. Now, anyone can predict the weather for tomorrow if he is broad enough in his predictions or if he is unclear in what it is that he predicts. For example: The weather tomorrow will be weather. It will be either rainy, sunny, or cloudy. It probably will not snow. Chances are the weather will be better than today.

What have I said? First, I can not miss on the first part of the prediction. From where I sit, I can be almost positive that it will not snow. I have covered most of the possibilities for weather, and any measure of rainfall or brightness of the sun would be an acceptable measure of the weather. Now, the last statement is a little more tricky. Better weather for tomorrow! What are the standards for better weather? I am a farmer and the crops need water. If it rains tomorrow and it did not rain today, the weather will be better tomorrow. I can measure this. If I am planning a picnic for tomorrow and if the sun shines brightly, it is better weather tomorrow. In other words, the weather, for someone, will be better tomorrow than it is today, and almost surely the temperature, rainfall or wind measures, hour-by-hour will not be identical to today's measures.

We plan a program of remedial English. Our objective is to improve the level of English achievement among our students. Have we said anything? What is improvement? What are our standards for measurement? How much is considered improved achievement? Who is to be improved? Are all pupils the same?

We have considered strategies for planning. We have been told that specific activities which involve time and resources lead us to an event. Other activities lead to other events, frequently simultaneously, which, when put together lead us to a goal. Education does not occur in a vacuum. Education does not limit itself to the classroom. Events which are unscheduled also lead us to achieve some objectives so that we must, therefore, be concerned with where we want to be as end product of education rather than measure the way we get to that objective or goal. If it is important that the pupil should use the English language properly, does it really matter whether or not he did his homework as it was assigned? Let us shift the emphasis upon measuring education to the objectives and goals of education rather than on the flurry of activities that lead to that goal. The spots on the chart of planning refer to events or happenings. The event must occur before we can move to the next event.

Having read the title of this presentation, and having been educated to the extent that you are able to "understand" the word objective, you are "aware" of education objectives. Having you return to today's session indicates you have "improved" over yesterday. You are a day older, aren't you? Isn't that better than not being able to return? These are not the real objectives, then, for my being here. The real objective is that you will be able to state an educational objective in terms of a measurable (or observable) criterion according to a specified standard. The subject of my objective is "you," the learner. The criterion which you will use as a measure are the inclusions of the four essential ingredients of the objective, upon which I shall elaborate. The standard will be self-imposed; in this case you will serve as the judge. The conditions are implied, in that you will perform this task either abstractly in your mind or concretely on paper before leaving this workshop today.

I am assuming that, having stated an objective in this manner, it will be an improvement over what has been an acceptable statement of an educational objective. My objectives, therefore, become the manifestation of the strategic positions you are to assume if you are to write an acceptable proposal, as a goal, and that you will be able to demonstrate or report that you have attained your objective at the termination of the grant period, or at least report the degree of attainment you have achieved. You will be able to proudly proclaim: "I did what I set out to do."

I mentioned that there are four elements of a good objective. The usual way of expressing a "good objective" is to say that it is in behavorial terms or is an operationally defined objective. This merely means that the objective expresses the behavior the learner will manifest when he has obtained the objective, or that it is how the system operates when the objective is met.

The first element, the learner, is the subject of the objective. Frequently, objectives, I think, improperly express what the teacher plans to do rather than how the pupil will behave. We in education have concern for the pupil and what he learns. It must be obvious, therefore, that the pupil must be the target for the objective.

Secondly, the verb for the objective must be an action that can be observed or measured. (In educational measurements, we usually call a measurement an observation.) Someone has said that all observable verbs can be classified into five distinct categories of verb forms

1. To identify
2. To name
3. To construct
4. To describe
5. To order

Obviously, the list omits to understand, to be aware of, to appreciate, to know. These latter verb forms defy observable actions because we can not know what takes place in a pupil's mind. We can observe choices of actions, repeated words or descriptions, and a stated order of importance.

Now for some details.

To Identify. The learner is expected to classify or to indicate membership or non-membership of a specified object or event into a class when the name of the class is given.

Ex. 1: Is this a red ball?
2: Was Woodrow Wilson a president of the United States?

To Name. The learner supplies the correct label for an object or event when the name is not given.

Ex. 1: What part of speech is the word "orange" in the sentence written on the board?
2: Who was president of the U.S. in 1920?

To Describe. The learner generates and reports the necessary categories of object properties, events, event properties or relationships relevant to a designated referent. Usually, the pupil has been previously instructed about what is an acceptable pattern for the description.

Ex. 1: Describe a square.
2: Describe the events in history that led up to the involvement of the U.S. in World War I.

To Construct. The learner generates a product which meets stated specifications. This may be a concrete product, such as a figure cut from a sheet of paper or a wooden structure, but it could be a product that is written, such as an essay or a picture that is drawn. An oral statement can also be a construction in the form of a speech or the singing of a song.

To Order. The learner arranges two or more referents in an order or sequence specified by the teacher. The ordering may also involve naming or identifying as well as describing. Sequential events or sequential objects are proper considerations.

Many tasks include more than one of the aforementioned verb forms, and whenever the verb is just as appropriate it seems ridiculous to make the substitution. For instance, "to translate" from one language to another may involve naming, constructing, and ordering, so it does not seem worth the effort to change the word used.

The next criterion is the standard or the quality of the objective product or behavior. In other words, how well do you anticipate the learner will be able to perform the task required? Frequently, the standard is expressed as a score to be attained on a test (or other criterion measure). The standard may also include a percentage of pupils of a particular class who are expected to attain the score. It is essential, of course, to specify in advance what the test situation will be. It is not sufficient to state as a goal that 90 per cent of the class will receive a grade of D or above because grading on this basis is arbitrary and controlled by the teacher. Not all teacher-made tests can specify that 70 per cent is passing. To digress, the object here is to determine how well teaching has progressed, not how poorly pupils learned.

The fourth criterion for a good objective is the statement of conditions. Under what circumstances will the objective be met? Do we expect a child to perform as well outside the classroom as in the class? Do we expect a child to run 100 yards as fast in a football uniform as in a track suit? Do we expect a pupil to recite words as readily from the printed page (reading aloud) as we do in normal conversational use of the same words?

It should be obvious by now that some of the criteria for educational objectives are implied rather than stated. The concern we have is for whom in the audience who is expected to read the objective. A spelling lesson may have a set of objectives for you as a teacher that implies that the child will write all of the words dictated by the teacher during a ten-minute spelling period and that each child will write legibly, etc. The same objective stated in a proposal for an experimental program, however, would specify the standards and conditions and the operation to be performed by the pupil.

Our next consideration should be focused upon when we write certain kinds of objectives. How complete should the objective be? The quick reply to the questions are in the form of other questions. How will the results be evaluated? What are the purposes of the evaluation? Who is to evaluate? To whom are you accountable? What decisions will be made when the objective is attained? What decision will be made if it becomes impractical to meet the objective?

In our strategy to meet an ultimate goal for our program, we establish enabling objectives. These objectives also form an hierarchy of objectives to be met as we proceed. We move from daily lesson plan objectives to project objectives. The daily objectives are very detailed and specific. The project objectives may be more general in nature, but each follows the pattern of the four elementary criteria. By the close of tomorrow's class I may state that each child in the class will be able to add 2 and 3 and arrive at the answer

5 when orally presented with the numerals two and three with instructions to "add." If it is presented as a role, drill. and practice exercise, the process may be "to name" the proper answer. If it is presented as a problem in logic, it may be a process "to order." The meeting of this objective at the termination of a particular unit (or time) is to be able to name the result of the process of adding two one-digit numerals when given two numerals at random. The objective for the end of the school year may be in the same form but may be a matter of naming the result when presented three one-digit numbers, etc. Ultimately, the goal for elementary school arithmetic may be stated in terms of being able to add no matter what the circumstances may be and the objective could be stated as a percentage of pupils at the end of the sixth year in school being able to score at a specified level on some standard achievement test.

Is the teacher actually expected to state objectives in this manner? Yes. Surely, the teacher has a purpose in mind for each day's learning session. Why not state the objectives to eliminate fuzzy teaching for nebulous goals that encourage procrastination? Secondly, if the objective is clear and has been definitely reached, why duplicate the effort just to be sure everyone has reached the goal? If the objective is not reached, it is time for a decision. Was the objective impossible to reach? Was the teaching ineffective? Do I re-double my effort to attain the same objective or do I re-state the objective? Is the objective essential in my strategy to reach the next objective?

One last thought. The behaviorally stated objective permits an evaluation of a rather specific situation. There is little point in evaluating on one basis when the teaching was toward another behavior. Meeting the criteria for a behaviorally stated objective leaves no doubt in the mind of the evaluator about what you intended to accomplish. Title III projects are experimental in nature. Future funding for projects will be based upon whether or not you are able to define what you have accomplished. In like manner, the continued support of public school is reaching a point when taxpayers are demanding a similar evaluation of success. Taxpayers are reluctant to support a school district that does not know what it is accomplishing.

Tests of knowledge which may be used as the criterion for success are, at their best, only indicators of what pupils know. Strive, therefore, to measure what a pupil knows in terms of what you expect him to know as it affects his behavior or actions. Avoid measurements that indicate what he does not know. The world is full of unknowns. We can teach the child only on the basis of building upon what he now knows. It matters not what you think you have taught; success is manifest only in what the child learns.

There are teachers who somehow have been led to believe that the purpose for stating objectives behaviorally is to ease the job of an evaluator. It is true that the evaluator's task is easier if the entire staff has well-stated objectives. The evaluator merely takes the data gathered on each criterion measured and reports the extent to which the objective is met. If this is true, objectives, perhaps, are of inestimable value to the evaluator. The evaluator has no more stake in the project than the teacher has—and frequently—not as much.

The very nature of the Title III project implies experimentation. Thus, it is implicit in the project implementation that new things are to be tried. The over-all effect of the project is the concern of the project backers (the money sources); but, surely, the desire of each teacher or other staff member is to produce a result that can be identified as a desirable practice worthy of attention. Specific objectives and specific feedback of the results are a part of the process so that proper alterations in the process can be made before harm results. Evaluation is concerned with the here and now cooperatively with the teacher. This is especially needed when new materials are being tried. The evaluator's research serves as the basis for change. At times, results are needed this week to identify what is needed next week. Without the able assistance of well-stated objectives, a teacher and an evaluator, working hand-in-hand, can do irreparable harm to the pupils in the project's care. Again, the objective is to describe where the pupil will be when the process has been completed. The objective is not written and data are not collected for the benefit of the evaluator.

The pressures on the project teacher are far greater than the typical classroom assignment. The responsibility frequently is a complete re-tooling which may throw out many years of accumulated lesson plans and procedures. Additionally, the pressure is upon the teacher to reveal to others what changes have been made in the innovative environment. Previously held, implied objectives for education are no longer valid. The specifics of the new objectives help to re-define goals for other teachers who may be expected to follow in the footsteps of the project teacher.

One tends to place importance upon what one sets forth in writing. The more precise the writing, the more precisely one tends to adhere to what has been stated. The great thinkers of history were those who were able to state clearly what it was they were thinking. The discipline of writing produced a carefully constructed, consistent statement of principles which was read and understood by those who followed.

6
AN ORGANIZATIONAL STRUCTURE OF VARIABLES AFFECTING EDUCATIONAL PROGRAMS

by
Terry D. Cornell

AN ORGANIZATIONAL STRUCTURE OF VARIABLES AFFECTING EDUCATIONAL PROGRAMS

Terry D. Cornell, *Coordinator*
Evaluation Services Division
EPIC Evaluation Center

Evaluation may be defined as a systematic procedure of collecting and analyzing information for the purpose of decision-making. The most important word that appears in this definition is probably "systematic." Systematic evaluation involves the process of very careful planning of the procedures which are going to be used to collect and analyze selected information.

One might conceptualize the systematic procedure in four stages. The first stage might be termed the Planning Stage or Planned Program. In this stage, the evaluator and those people who are involved in the educational program to undergo evaluation identify the variables which are affecting the educational program. Next, these variables are used to (1) write behavioral objectives, and (2) provide guidelines for the evaluation design that is to be used to evaluate these objectives. In addition, a program calendar is written containing four items: (1) the behavioral objectives of the program, (2) the program description which involves a very detailed description of the variables which have been identified as affecting or interacting in this given educational program, (3) a monitoring system which can be used to collect information for the purpose of determining whether or not what is planned is actually what is implemented in the next phase, and (4) a calendar of events that indicates the dates at which the data are going to be collected and by whom.

The second stage might be called the Implementation Stage or Actual Program. This is the point in time where the planned program described in the first stage is actually implemented. At this time, the monitoring system is applied to collect information on the actual program activities. This information is then compared to what was planned in Stage I to see if everything is going as planned. Due to this comparison, certain changes might have to be made in the evaluation design or in the objectives to represent what is actually being implemented.

Output of the program is dealt with in the third and fourth stages. Stage III might be termed the Feedback Stage. At this time, the data which are collected in Stage II are analyzed and subjected to statistical tests. The results, then, are used to make decisions as to whether or not the objectives stated in Stage I were met.

Stage IV is the Recycling Stage. At this point, the program is reinitiated into the first stage for the consideration of those objectives which were not met and also to consider new variables, and in turn, new objectives, that might be evaluated in the next cycle. The Recycling Stage implies that evaluation is a never-ending process and may continue for many years until all variables which are affecting a given educational program have undergone evaluation.

To ensure that systematic evaluation takes place, one must first realize that any educational program is affected by many variables, not only within the educational setting, but also outside of it. These variables may combine in many different ways to produce different kinds of outcomes in the education program. It would, therefore, be helpful to the evaluator if these variables could be displayed in such a way that he could visually see the different ways in which the variables could come together to affect the outcomes of an educational program.

The Structure of Variables is a means of identifying and displaying the elements which describe the educational program and its related variables. Three dimensions are utilized with each of the variables assigned to one of these dimensions. In the Tennessee Design for Evaluation these dimensions are Program, Population, and Behavior. At EPIC, these are called Instruction, Institution, and Behavior.

INSTRUCTIONAL DIMENSION

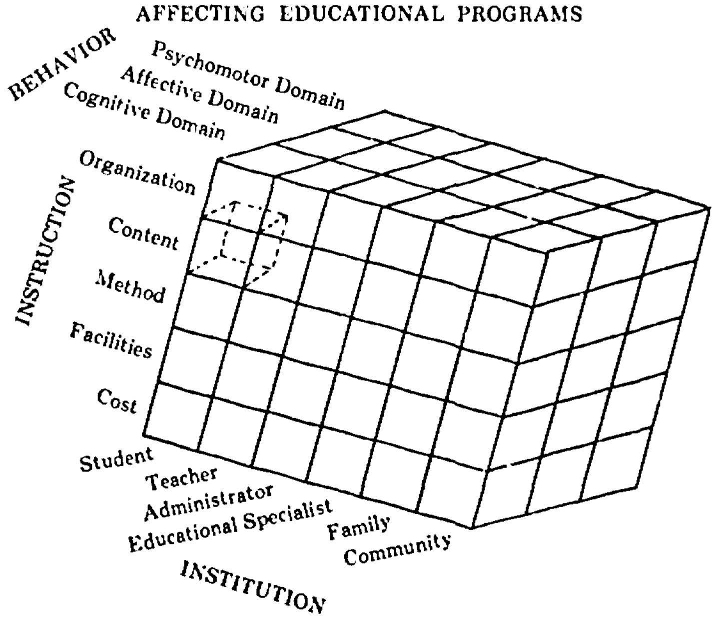
The Instructional Dimension is that dimension of the model which describes the innovation in terms of specific variables. (This is identified as the Program Dimension in The Tennessee Design). The first of these variables is Organization which is defined as the matrix in which teachers and pupils are brought together so that instruction can take place. The organizational matrix may be divided into two components known as time and space.

1. Time refers to the duration and sequence of blocks of time devoted to the subjects taught. Duration may be defined as the length of any given period. Sequence may be defined as the order in which subjects are taught. Duration and sequence may be thought of in terms of both daily and weekly scheduling. (Example: Science may be taught only twice a week.)
2. Space refers to the vertical and horizontal organization of students. Vertical organization serves to classify students and move them upward from the point of admission to the point of departure. Horizontal organization divides students among teachers. Both grouping processes may be homogeneous, heterogeneous, or a combination of the two.
   a. **Vertical Organization**: Vertically, schools may be graded or non-graded, or fall somewhere in between.
      (1) Graded: In pure grading, the content of the instruction program and its sequential arrangement are determined by assignment of subject matter to various grade levels, by designation of instructional materials suitable for particular grade levels, and by promotion of pupils upon satisfactory completion of the work specified for each grade.
      (2) Non-graded: In pure non-grading, sequence of content is determined by the inherent difficulties of subject matter in

the childrens' demonstrated ability to cope with it; materials are selected to match the spread of individual differences existing within the instructional gap; and the children will operate according to their readiness to perceive. Promotion or non-promotion does not exist as such. An important goal is to provide for the continuous progress of each child.

Figure 3

AN ORGANIZATIONAL STRUCTURE OF VARIABLES AFFECTING EDUCATIONAL PROGRAMS



b. **Horizontal Organization:** Horizontally, schools may be organized into any one of many alternative patterns. But all of these horizontal patterns are derived from essentially four different kinds of considerations—considerations of the child, of the curriculum, of the teacher's qualifications, and of the school's philosophy.
   (1) Self-contained: A self-contained classroom is defined as a classroom in which a group of children of similar social maturity, ability, age, etc., are grouped together under the continued guidance of a single teacher.
   (2) Departmentalization: The characteristic feature of departmental instruction is that a teacher who is highly trained in a field of knowledge is assigned to teach that academic subject.

(3) Cooperative teaching: Under the general heading of cooperative teaching may be found dozens of different patterns of school and staff organization. Some of these are derived from, or associated with, attempts to achieve greater flexibility in pupil grouping. Others are associated with efforts to eliminate the administrative and instructional characteristics of rigid, lock step, organizational structure. One of the most important forms of cooperative teaching is the organizational pattern known as team teaching.

The second variable is Content. Content is defined as that structure or body of knowledge which is identified with the subject matter of a discipline and controls its inquiries. Content may be described in terms of specific topics to be covered at a given grade level.

A third variable is Methodology. Methodology is that process designed to facilitate learning. It may be divided into three levels: teaching activities, types of interaction, and learning principles or theories utilized.

1. **Teaching Activities**
   a. Lecture
   b. Discussion
   c. Question-answer
   d. Committee
   e. Round table
   f. Symposium
   g. Drill
   h. Homework
   i. Review
   j. Individual supervised study
   k. Resource person(s)
   l. Field trips
   m. Inquiry
   n. Debate
   o. Media*

*Includes texts, resource books (dictionaries, encyclopedias, library, etc.), workbooks, films, filmstrips (with and without tapes), tapes/records, television (commercial, educational, closed-circuit), laboratories (science, language), programmed teaching machines/texts.

2. **Types of Interaction**
   a. Teacher ⟷ Student
   b. Student ⟷ Student
   c. Media ⟷ Student
   d. Teacher ⟷ Teacher*

*Principally team teaching. (In addition to identifying the interaction participants, there are a number of codes that have been developed to describe the interaction, such as: (1) Interaction Analysis—Ned Flanders; (2) Teaching Interaction—Marie Hughes; (3) Classroom Transaction—Stanford University.)

3. **Learning Theory**
   a. Behavior which represents the achievement or partial achievement of an educational objective should be reinforced.
   b. The introduction of cues which arouse motivation toward the achievement of an educational objective will increase the effectiveness with which that objective is achieved.

c. Practice in applying a principle to the solution of problems will increase the probability of transfer of training to new problems requiring the use of the same principle for their solution.
d. Since learners differ in their capacity to make the responses to be acquired, learning will be most effective if it is planned so that each learner embarks on a program commensurate with his capacity to acquire new responses.
e. If a pupil has had training in imitation, then he is capable of learning by observing demonstrations of skills to be acquired.
f. The learner will learn more efficiently if he makes the responses to be learned than if he learns by observing another make the responses or makes some related response.

The fourth and fifth variables are Facilities and Cost. Facilities is defined as that space, special equipment, and expendables needed to support an educational program. Cost is the money required for facilities, maintenance, and personnel to accomplish a given task.

The variables defined in the above represent important categories to be considered in the educational program. The innovation to be considered may be contained in any one of the variables (e.g., team teaching—organization). Yet all variables must be considered in the analysis of the total program. If innovations are to be adopted on a wide scale, a complete picture of the program must be studied with its various components carefully analyzed.

INSTITUTIONAL DIMENSION

The Institutional Dimension is that dimension of the model defined by the variables of Student, Teacher, Administrator, Educational Specialist, Family, and Community. Any given innovation will be influenced by the unique qualities of the individuals involved. For the purposes of evaluation, each of the variables is described in terms of sub-variables that may have a direct influence on the given program. The following examples are a sample of these descriptive sub-variables:

1. **Student**
   - a. Age
   - b. Grade level
   - c. Mental health
   - d. Sex
   - e. Familial variables
   - f. Socio-economic variables
   - g. Physical health
   - h. Achievement
   - i. Ability
   - j. Interest
   - k. Relationship to program
2. **Teacher, Administrator, and Education Specialist**
   - a. **Identification Data**
     - (1) Age
     - (2) Sex
     - (3) Race, nationality, religion
     - (4) Physical health
     - (5) Personality characteristics
   - b. **Educational Background and Work Experience**
     - (1) Undergraduate major and minor
     - (2) Graduate major
     - (3) Highest degree
     - (4) Educational experience
     - (5) Experience outside Education

c. Environmental Factors
   (1) Professional salary
   (2) Professional affiliations
   (3) Non-professional affiliations
   (4) Socio economic status of residence
   (5) Professional and non-professional reading habits
   (6) Leisure activities outside professional work time
d. Degree of Involvement in Program

3. Family
   a. Degree of Involvement with Program
      (1) Have children in school; all affected by the program.
      (2) Have children in school; some affected by, some not affected by, the program.
      (3) Have children in school; none affected by the program.
      (4) Have no children in school (these are treated under descriptive items in the Community variable).
   b. General Characteristics
      (1) Ethnic/national/linguistic
      (2) Size
         (a) Total
         (b) Siblings
         (c) Other relatives present
      (3) Age distribution
      (4) Marital status
      (5) Pattern
         (a) Nuclear
         (b) Extended
      (6) Income
         (a) Approximate level
         (b) Number of wage earners
         (c) Source
         (d) Occupation
      (7) Residence
         (a) Urban
         (b) Suburban
         (c) Rural
         (d) Cost range
      (8) Education
         (a) Approximate formal level
            i. Parents
            ii. Siblings
            iii. Other relatives present
         (b) Informal
            i. Industrial
            ii. Military
            iii. Community Service
            iv. Other
      (9) Affiliations
         (a) Religious
         (b) Political

(c) Social
(d) Professional
(e) Other
(10) Mobility
(a) Parents' place of origin
(b) Length of time in community
(c) Frequency of moving
(d) Extent of traveling

4. Community
a. Geographical Setting
(1) Location
(2) Environment—general
b. Historical Development
c. Population Characteristics
(1) Demographic data
(a) Population size
(b) Population density
(c) Marriage and divorce rates
(d) Birth and death rates
(e) Age distribution
(2) Ethnic/nationality
(3) Linguistic
(4) Change patterns
(a) Mobility patterns
i. Immigration
ii. Emigration
iii. Migrant-indigenous ratio
(b) Growth patterns
d. Economic Characteristics
(1) Commercial/industrial organization and development
(2) Occupational range
(3) Sources/range of individual incomes
(4) Sources/range of tax base
e. Social Characteristics
(1) Institutions and organizations
(a) Government/political
(b) Educational
(c) Religious
(d) Service
(e) Social
(f) Commercial/financial
(g) Labor
(h) Professional
(i) Recreational
(j) Protection
(2) Power structure
(3) Socio-economic stratification

Assessment programs of the past have focused primarily on the child and his response to content in a given subject area. With the changes taking place in educational programs, more evidence is needed about the influence of the teacher, administrator, parent, and community on a given program.

BEHAVIORAL DIMENSION

The Behavioral Dimension is defined by the variables of Cognitive, Affective, and Psychomotor Behavior. Cognitive Behavior includes the recall, comprehension, and application of knowledge and the utilization of intellectual skills of analysis, synthesis, and evaluation.

The second variable in this dimension is Affective Behavior. Affective Behavior is defined as the interest, attitudes, values, appreciations and adjustments of the individual.

Psychomotor Behavior is the third variable in this dimension. It includes those acts which involve neuro-muscular coordination. Handwriting and physical education utilize this variable to draw conclusions about special programs.

In summary, the Organizational Structure of Variables Influencing Educational Programs provides a framework to identify the forces affecting a given program. The factors created by the interaction of one variable from each of the dimensions may be studied in depth by any school district.

The study of these variable interactions becomes, then, the evaluation dimension of an educational program. The question then becomes, "How might the Organizational Structure of Variables facilitate the evaluation of an educational program?" As was stated previously, the description of the planned program includes the identification of those variables that are felt to be affecting the educational program under consideration. These variables will depend upon the proportion of that educational program selected for study. After the variables have been identified, they can be combined into factors. A factor is simply a combination of one variable from each of the three dimensions. For example, if the planned program under consideration was a study of the student's achievement in math, the factor from Figure 3 that would be defined would be: Student-Content-Cognitive Behavior. If one was interested in the study of a student's attitude toward mathematics, the factor identified would be: Student-Content-Affective Behavior (refer to Figure 3). Once the factors have been formed, the behavioral objectives can be written. Essentially, four questions might be answered by a behavioral objective:

1. What is the Institutional Variable?
2. What is the Instructional Variable?
3. What is the Behavioral Variable?
4. How is the behavior going to be measured?

The task of writing objectives becomes a very easy one, since three of the four questions have been answered as a result of the factors produced through the identification of the variables influencing the program. An example of an objective which answers these four questions is: "When exposed to the individualized reading program, the student will improve his reading comprehension skills as measured by the Silvaroli Reading Inventory."

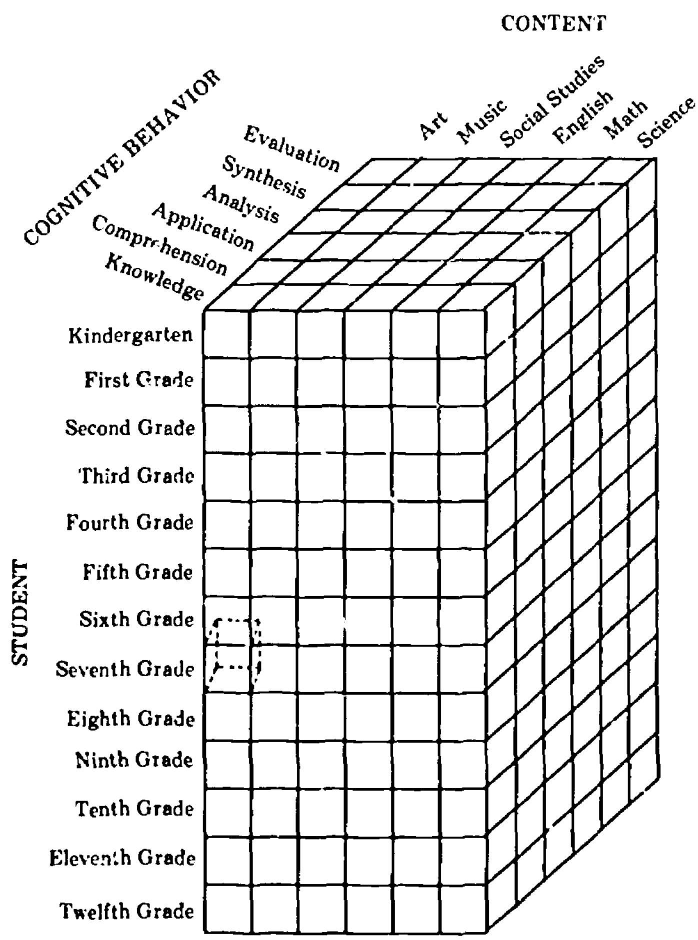
In analyzing this objective, the Institutional variable is Student, the Instructional variable is Content, the Behavioral variable is Cognitive Behavior—Comprehension Level, and the measurement is the Silvaroli Reading Inventory.

If the Organizational Structure of Variables Influencing Educational Programs is utilized in writing behavioral objectives, it is not necessary to use one system to write objectives at a program level and use another system, such as Mager, to write objectives at the instructional level. In other words, the level of specificity in writing behavioral objectives depends upon the level of specificity used in describing the variables of the educational program to undergo evaluation. For example, utilizing the original structure displayed in Figure 3, an objective might be: "The student will gain cognitive skills in selected curriculum areas as measured by pre-determined tests." This objective answers the four questions and will be termed a behavioral objective. However, taking that factor, Student-Content-Cognitive Behavior (refer to Figure 3), it can be broken down into a number of sub-variables. For example, the Student variable could be broken down in terms of students, kindergarten-twelfth grades. The Content variable could be broken down according to selected subject matter areas in a given curriculum, such as math, science, English, social studies, etc. The Cognitive Behavior variable could be broken down according to the levels of Bloom's *Taxonomy*. Consequently, as indicated in Figure 4, the factor, Student-Content-Cognitive Behavior, has now become a new Organizational Structure of Variables, but more specific in nature. The resulting objectives using this structure will also be more specific. If we selected the sub-variables of fifth grade student, science, and knowledge to undergo evaluation, an objective that might result would be: "The fifth grade student will gain knowledge in science as measured by the Stanford Achievement Test." Notice that this objective, when compared to the first one, is much more specific. This is due to the fact that the variables we are using are much more specific when you compare them to the variables of the original Structure.

Again, we could become more specific. Figure 5 shows an even more specific structure when taking the factor, Fifth Grade Student-Science-Knowledge. The fifth grade student sub-variable could be subdivided even further in terms of grade placement levels. Science could be broken down into topics, such as physics, chemistry, general science, biology, botany. Knowledge could be broken down even further according to Bloom into knowledge of specifics, knowledge of ways and means, and knowledge of universals, A resulting objective might be: "A fifth grade student with a grade placement of 5.5 will increase his knowledge of specifics in biology as measured by questions 1, 3, and 5 found on the Stanford Achievement Science Sub-test." Again, we have become more specific in our variable description and consequently in our objective. So, in a sense, then, the same procedure for writing objectives can be utilized to write a very general program objective as well as a very specific instructional objective. The only difference in the objectives is that one is more specific than the other because the variables are much more specific. However, no matter at what level of specificity objectives are written, a structure of variables derived from the original Structure is utilized. One can readily see that the Organizational Structure of Variables Influencing Educational Programs really has an infinite number

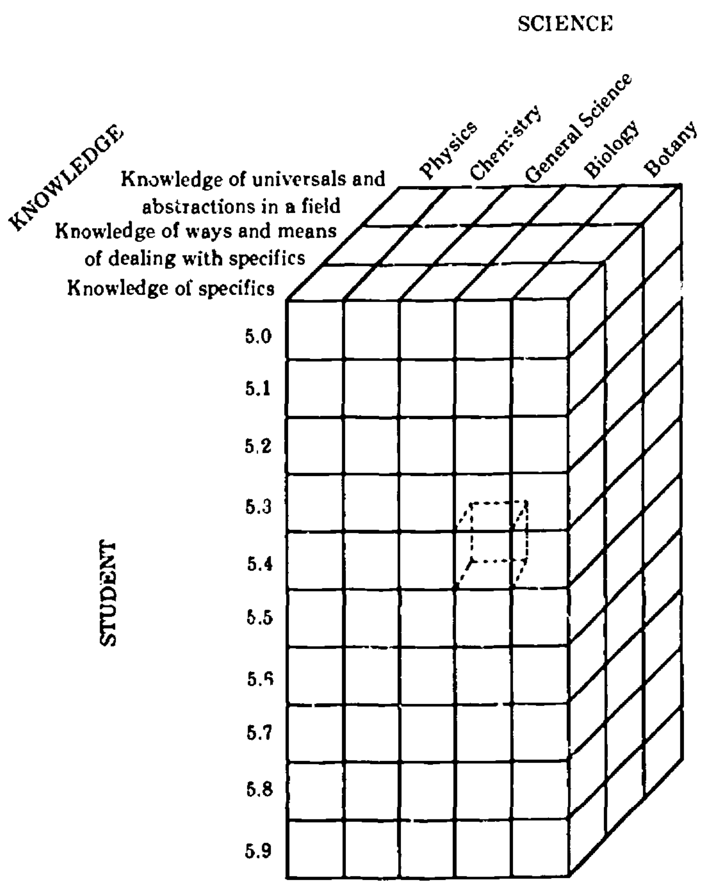
Figure 4

STRUCTURE OF VARIABLES: CONTENT



72

Figure 5
STRUCTURE OF VARIABLES: SPECIFIC CONTENT OF SCIENCE



of levels of specificity and the given level of specificity that an educator uses to write the behavioral objectives of an educational program will depend on the amount of time he intends to use in accomplishing the objectives. In other words, as the degree of specificity increases the amount of time needed to accomplish the objective tends to decrease.

Basically, then, at this point we have identified two ways in which the Organizational Structure of Variables can be used. It can be used to identify those factors which are influencing an educational program. Secondly, the Structure can be used as a system for writing behavioral objectives. Another way that the Structure could be used is to act as a storage system for educational programs, objectives, tests, and specific test items. By this is meant that instructional programs that have undergone evaluation or test items that have been used in the evaluation of a given behavioral objective can be classified according to the factor in the Structure identified for that program or test item. In other words, an item measuring math achievement for a student would be placed in the Student-Content-Cognitive Behavior factor. An instructional program dealing with changing a teacher's attitude toward the methods that she is using would be categorized Teacher-Method-Affective Behavior. So, in a sense, the Organizational Structure becomes a filing system for keeping a record of tests, objectives, and/or programs which have undergone, or have been used in, evaluation. Utilizing a computer system, the Structure could become a very effective means of storing information.

The Structure can also be used as a basis of communication among educators. The definitions of the variables in the Structure become a basis of communication among educators, and the problem of semantics and the many connotations connected with these terms are then eliminated and communication is facilitated.

Another way that the Structure might be used is in curriculum development. If the activities and the programs in a school district are classified or categorized using the Structure, one might find that there is a real lack of instructional programs or curriculum development in given areas of the Structure. This, then, would make teachers and administrators aware of the fact that work needs to be done in a given area or areas.

In summary, then, the Organizational Structure of Variables can be used to provide for the systematic procedure of evaluating an educational program by giving the evaluator guidelines in which to identify those variables influencing a given program and also providing him with definitions of variables which he can then incorporate into behavioral objectives. Also, the Structure provides him with information which might lead to the development of educational materials and techniques that would undergo evaluation at a future time.

Although the above is important, probably the most important use that the Organizational Structure of Variables has is making educators aware of the many variables that affect educational programs and consequently the complexity of educational evaluation.

7
MEASUREMENT INSTRUMENTS

by
Willis C. Nowell

MEASUREMENT INSTRUMENTS

Willis C. Nowell, *Director*
School Psychological Services
Tennessee State Department of Education

INTRODUCTION

In the time alloted, we are not able to trace the chronology of the development of measurement in education. Yet, it is interesting to note that Isaac Daughton in 1935 wrote: "Whatever exists at all exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality." In 1923 William McCall wrote: "Measurement is indispensable to the growth of scientific education . . . the final answer to every educational question, except one, must be left to the educational measurer and must await the development of education as a science." These statements are quoted by Ross and Stanley[1] in their discussion of the state of education as an art or science. One thing that emerges from a review of the literature of the early development of educational measurement is that any attempt to quantify achievement through instrumentation is today, as it has always been, suspect.

The Gestalt psychologists point out: "Any organism is more than its parts." Therefore, the Gestaltists generalize that adding test scores does not give you a valid measure of the total person. And yet, they conclude that a reading score tends to measure the entire organism functioning in that reading situation.

Kilpatrick points out that in the decision-making process, where input is important, verification is not an appropriate term and techniques of measurement are not in themselves adequate. He adds, however, in a conciliatory manner, "In such cases the function of measurement is not to supplant or to supply decisions, but to furnish, regarding the working of the policy under review, more and better data, in the light of which a fresh and better decision can be made."

That is precisely the dimension measurement specialists are attempting to add to the educational process. Evaluators concern themselves with collecting objective data that can be used as input information to aid in the decision-making process regarding the development of a learning climate that is conducive to cognitive and effective growth. It should, however, be remembered that measurement "is not an end, rather it is a means to an end." Measurement is generally a qualitative description of something that is observed. Measurement data, taken out of context, is not very meaningful. Such data becomes meaningful only when reviewed in concert with relevant facts and considered as a part of the total constellation of stimuli.

We need not, however, defend or present an argument at this time for the use of measuring devices and for appraising the relative effectiveness of

[1]Most of the principles enunciated in this paper were extracted from the third edition of *Measurement In Today's Schools* by C. C. Ross and Julian C. Stanley.

treatment techniques employed by modern curriculum designers. Psychologists and educators have already fought—and to some extent won—that battle for us.

It is our purpose here today to describe the pattern of measurement currently being employed in public education and to discuss the contemporary tools used in measurement. Ross and Stanley categorize and classify measuring instruments into several categories. They present oral and written tests and contrast them. They proceed to divide written tests into non-standardized, informal essay, and objective types. The formal written tests discussed by these authors are what scholars commonly refer to as standardized tests. Under formal standardized tests, instruments have been developed to measure achievement, including the general survey type instrument and specific diagnostic instruments for measuring and analyzing achievement. Standardized intelligence tests are also classified as formal tests and are divided into general, individual, and group tests as well as specific aptitude or prognostic tests. And finally, the authors show personality and interest tests as a form of standardized instruments used in the total evaluation of the individual.

They point out that the distinction between oral and written tests is obvious. Neither is it difficult to distinguish between informal and formal tests.

On the other hand, the distinction between tests of achievement, intelligence, personality, and interest are not so easy to make. For this reason, it may be important to define briefly what is meant by tests named in the above categories.

Achievement tests, for example, are exactly what is implied by the term itself, that is, instruments designed to measure the academic achievement of subjects. Personality and interest tests are self-report type instruments rather than measures of ability. Intelligence tests are generally defined as measures of learning capacity as opposed to achievement tests which attempt to measure learning itself. It should be noted, however, that most intelligence tests attempt to measure educability while achievement tests are usually limited to the measuring of education.

One may conclude from these definitions that achievement and intelligence are aspects of personality. For personality is employed by psychologists to include every trait that differentiates one individual from another. If that is true, then personality may be measured by all types of tests, including rating scales, questionnaires, interviews, and controlled observations.

Tests are generally subdivided into three categories: general and specific on the basis of scope, individual and group on the basis of administration, verbal and non-verbal (or performance) on the basis of content.

It should be noted that most standardized instruments are tests rather than scales. Only on rare occasions are scales subjected to standardization processes. Test items arranged in order of difficulty usually constitute a scaled test.

TEST CHARACTERISTICS

With this frame of reference, we may now appropriately ask the question, What are the characteristics of a "good" test? There are certain essential characteristics of a dependable test, whether it is a commercialized standardized test or a teacher-made test. These characteristics are described semantically by Ross and Stanley as validity, reliability, and usability.

Teachers have an obligation to understand the meaning of these terms and, certainly, a prerequisite for evaluation specialists employed in projects, such as the Elementary and Secondary Education Act, Title III, Demonstration Programs is a clear familiarity with the meaning of measurement terminology.

Validity is generally defined as a degree to which the test or instrument measures what it claims to measure. That is the truthfulness of the instrument. For example, in arithmetic reasoning tests, it is important to know whether the instrument is actually measuring reasoning ability, arithmetic problem solving ability, or something else, such as reading ability or general intelligence. Very often it is difficult to determine what the instrument is actually measuring. This characteristic of the test, validity, is the most important component of any measuring device. Test developers and users then must preoccupy themselves with determining the validity of the instrument. There are accepted methods of doing this. In the first place, a very clear statement of the specific objectives of the course or subject must be defined. You must know what it is you wish to measure. In the second place, the determination of the standardization sample must be made and the standardization sample must be adequate. And finally, the accuracy of the instrument itself must be determined. The precision of discrimination of each item is most important.

Validity implies specificity and should always be considered in that light. Tests are valid for one purpose. Users of tests should examine the test manual before determining if the instrument is suitable for their purpose. In most cases, the test manual will show who developed the test, how the items were selected, what validation and standardization procedures were followed and other pertinent information vital in the selection of a test.

In addition to these data, one should examine the content of a standardized test. Many are invalidated with errors. In some tests with as many as three-thousand items, there are three to four hundred errors easily detected upon close scrutiny. A test that is not well standardized with clearly stated validity and reliability coefficients should be rejected by the user. In cases where measurement specialists are not available for consultation with local school officials, it is best to consult statisticians, psychologists, or other measurement specialists who may be located at large university research centers. To make reference to such books as Buros' *Mental Measurement Yearbook* may be helpful.

Validity has been discussed at some length and it is now important to refer to reliability as a term that lends meaning to the value of a test. Reliability refers to the degree to which a test agrees with itself or the extent that two or more forms of the same test will give the same results when repeated. Reliability, then, simply means consistency. Reliability differs from validity in that reliability has nothing to do with truthfulness. Although validity is the most important attribute of the test, reliability is certainly the second most valuable characteristic. One may conclude, then, that the ideal test is the test that tells the truth consistently.

It is not a good practice to study reliability values of a test without considering validity. It should be noted that reliability does not ensure validity.

As with validity, there are a number of ways to determine the reliability of a test. The multiple form technique is probably the most common method utilized by test builders. In such cases, more than one parallel form of the test is developed. These forms are administered to large numbers of subjects with the expectation that the scores obtained for the different forms will agree very closely. Another method of determining the reliability of a test is what is commonly called the test-retest method. This technique involves the initial administration of the test and after a relatively long period of time repeating the same form of the test, computing the coefficient of correlation between the two series of scores to determine if the findings are essentially the same. And finally, a third method is called the split-half technique. This procedure involves the division of the test into halves, matched for content and difficulty. When the two series of scores are obtained, the coefficient of correlation is computed.

It should be noted that this technique gives only the reliability for the half-test. A very difficult formula is involved in determining the reliability from this technique of the total test. Although reliability can be determined, there is no clear way to state the critical point of reliability. It simply should be remembered that a test should be as reliable as possible for the purpose for which it was intended.

Ross and Stanley, finally, point out that usability is an important characteristic of a test. Usability, as defined by these authors, is the degree to which the test or instrument can be successfully employed by classroom teachers and school administrators without an undue expenditure of time and energy. There are at least five criteria for usability. A test should be easy to administer, easy to score, easy to interpret, cost very little, and be mechanically sound.

It is convenient for educators to simply order standardized tests developed by commercial organizations to measure academic achievement, intelligence and personality. There are cases currently being encountered quite often—in fact, since the introduction of federal treatment programs—where no standardized tests, commercially prepared, are available to measure the characteristics desired for decision-making. In such cases there should be no reason why a teacher could not construct a test that would serve his purpose.

TEACHER-MADE TESTS

There are several reasons why teachers should learn to build tests. Most tests used in classrooms are teacher-made, and it has been demonstrated that essay type tests most commonly used are not very valid and reliable. Yet, skillfully prepared, informal tests are as reliable and valid as some standardized tests. A teacher-made test may be logically prepared through the following processes: a careful planning process, preparation of items to be used in the test, testing the instrument on a group of subjects, and, finally, evaluating the test's effectiveness.

When planning the test, the teacher must consider the nature of the objective to be measured, the purpose it is to serve, and the conditions under which it will be used. The teacher or committee, which should always be formed to help develop a test, should evaluate all the important outcomes of instruction, including the development of an educational philosophy for the school. Objectives should be identified locally, although it is helpful to make reference to such sources as Bloom's *taxonomy of objectives*, or Rath's *Evaluating the Program of a School*, or some other test that would help to guide the committee in formulating operational objectives. Rath, for example, states several objectives, including functional information, various aspects of thinking, attitudes, interests, aims, purposes, appreciations, study skills, work habits, social adjustments, social sensitivity, creativeness, and a social philosophy.

When preparing the test, a preliminary draft of the test should be prepared as early as possible. The test should include more than one item. Most of the items in the final test should be of approximately 50 per cent difficulty (at least one-half of the subjects should know the answers to the items). The test should include more items in the preliminary draft than will be needed in the final draft, since revision of the test after a lapse of some period will result in the elimination of certain items that do not discriminate. The test items should be phrased in such a manner that the content rather than the form of the statement will determine the answer a subject will provide. The items should be worded so that the whole content functions in determining the answer rather than only a part of the test. All the items of a particular type should be placed together in the test. The items should be arranged in ascending order of difficulty with a regular sequence in the pattern of responses being avoided. Provision should be made for a convenient written record of the pupil's responses and, finally, the directions to the pupil should be as clear, complete, and concise as possible.

Beyond this step, the teacher or committee charged with the responsibility of developing a test should try out the test. This procedure is carried out to determine or locate the weak items and to identify the finely discriminating items that will remain as a part of the test. Four principles are usually involved in trying out the test, including allowing a generous amount of time for the administration of the test, a scoring procedure that is simple, and the preparation of answer keys prior to the actual scoring of the test.

The final step in the development of a teacher-made test is the evaluation of the test itself. This procedure is conducted to determine if the test is valuable or to determine if the test should be discarded. One may conclude roughly that the difficulty of the test offers some indication of its adequacy, that internal consistency of the individual items is determined by their ability to discriminate between pupils who rank high and those who rank low on the test as a whole. It is also a good practice to have the items interpreted or criticized by persons who have taken the test, and test results should be checked against an outside criteria, such as the ranking of students by a teacher who is familiar with their abilities. Finally, it is desirable to obtain the reliability coefficient of the test.

SUMMARY

It is clear from this brief discussion of measurement, with some reference to the mechanics of test construction, that the selection of standardized tests or the building of a new test is a very difficult task. It is important to reflect upon a quotation from Scates before we discard the idea of the importance of measurement. Scates asked:

> What has the measurement movement done for us? One way of responding to this question is to note what we would not have if we had no such measurements. Certainly we would not have much of our current scientific work in education for there cannot be a science without fairly precise quantification. Not that science is measurement but that traits which are devoid of any reasonably definite quality simply do not have the required specificity for entering into the careful thinking essential to science. When quantities are disregarded, almost any generalization becomes true. There is an infinitesimal element of truth in practically anything which might be said. Quantities are part of the nature of truth and are therefore an essential part of science.

8
APPLICATION OF ANALYTIC TECHNIQUES

by
Harry Bowman

APPLICATION OF ANALYTIC TECHNIQUES

Harry Bowman, *Research and Evaluation Assistant*
Southeast Educational Laboratory
Atlanta, Georgia

This presentation is organized around a discussion of some terminology, appropriate designs for school research-evaluation, and selected procedures for analyzing data. Although a brief review of terms may be unnecessary for some persons in this group, it is necessary to establish a common understanding before moving on to the other topics.

Measures of central tendency for a set of scores include the mean, median, and mode.[1] The mean is the arithmetic average of a set of numbers or scores. The mean is calculated by adding the scores and dividing the sum by the number of scores. For example, the mean for scores of 50, 48, and 43 is 47 (141 divided by 3).

The median of a set of scores is a value that divides the set into two groups with half the scores being larger and half smaller than the value. The median for scores of 9, 10, 12, 13 and 15 is 12 since two scores exceed 12 and two scores are less than 12. For a set of scores with an even number of scores, the median would be the average of the two scores in the middle of the ranked distribution.

The mode is an infrequently used index of central tendency. It is defined as the most frequently occurring score in a set of scores. The mode for the set containing scores of 27, 28, 29, 29, and 30 is 29 because that score appears twice and all other scores appear only once.

Measures of variability provide information about the spread of a set of scores around the center of a distribution. These measures include range, interquartile range, and standard deviation.[2]

The range of a group of scores is the difference between the largest score and the smallest score. The range for a group of scores of 15,18, 19, 21, and 25 is 10 because the largest score is 25 and the smallest score is 15.

The interquartile range of a set of scores is defined as the difference between the raw scores corresponding to the 75th and 25th percentiles. (The 75th percentile is the score which exceeds 75 percent of the total number of scores in the set.) For example, on the Metropolitan Readiness Test (1965 edition), a total raw score of 67 has a percentile rank of 75. A total raw score of 41 has a percentile rank of 25. The interquartile range for this test is 26 (67-41).

---

[1]Helen M. Walker and Joseph Lev, *Elementary Statistical Methods* (rev. ed.; New York: Holt, Rinehart, and Winston, 1958); and W. James Popham, *Educational Statistics: Use and Interpretation* (New York: Harper and Row, 1967).

[2]*Ibid.*

The standard deviation is a measure of variability based on the average squared deviation of the scores from the mean for the group. (The variance is the square of the standard deviation.) The calculation of the standard deviation, designated by s, may be accomplished with two formulas:

$$(1)\quad s = \sqrt{\frac{\Sigma x^2}{N-1}}$$

where x = a score (X) — the group mean ($\overline{X}$)
Σx^2 = sum of squared x for each score
N = number of scores

$$(2)\quad s = \sqrt{\frac{\Sigma X^2 - (\Sigma X)^2}{N-1}}$$

Where X = score
ΣX = sum of scores
X^2 = squared score
ΣX^2 = sum of squared scores
N = number of scores

The two procedures for computing the standard deviation are illustrated by the following example:

Score		Mean		x	x^2	X^2
13	–	8	–	5	25	169
11	–	8	–	3	9	121
10	–	8	–	2	4	100
7	–	8	–	–1	1	49
4	–	8	–	–4	16	16
3	–	8	–	–5	25	9

$\Sigma X = 48$
$N = 6$
$\overline{X} = 8$
$\Sigma x^2 = 80$
$\Sigma X^2 = 464$

$$(1)\quad s = \sqrt{\frac{\Sigma x^2}{N-1}} = \sqrt{\frac{80}{6-1}} = \sqrt{\frac{80}{5}} = \sqrt{16} = 4$$

$$(2)\quad s = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}} = \sqrt{\frac{464 - \frac{48^2}{6}}{6-1}} = \sqrt{\frac{464-384}{5}}$$

$$= \sqrt{\frac{80}{5}} = \sqrt{16} = 4$$

Of course, the second procedure in the example is the preferred way to compute a standard deviation because the mean, which must be computed to use the first procedure, is usually not a whole number and the second procedure is easier to use with large groups.

Other terminology that is frequently used in connection with research design includes random assignment, treatment-experimental group, control group, and measurement. Random assignment means that any available individual is as likely to be selected as any other individual for assignment to a group. This may be accomplished by assigning numbers to all the individuals, placing the numbers on pieces of paper, and drawing the numbers from a container.

The treatment-experimental group refers to the individuals who receive special instructional methods or materials designed to produce some change or improved learning. The control group is a group comparable to the treatment group but does not receive the special methods or materials.

Measurement refers to the assignment of numbers to events according to some rule or a qualitative description of something observed. Measurement instruments include a variety of procedures, such as tests, questionnaires, rating scales, checklists, and tabulation of events. Measurement is a basic part of the assessment of project or program success and requires careful consideration before implementation of any project or new program.

Data analysis for project evaluation cannot be viewed meaningfully except in relation to the research or evaluation design that applies to the project. Consequently, at this point, attention will be focused on a brief discussion of some selected research designs.

The first design may be referred to as the Nonequivalent Control Group Design.[3] This design is used in a situation where a pretest and posttest are used but individuals cannot be randomly assigned to groups. The design may be illustrated as follows:

```
           R
M ———————> T ———————> M
.......................
M ——————————————————> M
```

M = Measurement
T = Treatment
R = Random assignment

The Nonequivalent Control Group Design would be used when a treatment is introduced after class groups have been formed. While the groups must remain intact, the random assignment of the treatment may be made.

The design cannot be used if the groups are markedly different on a characteristic that is significantly related to the treatment. A pretest is

[3]Richard E. Schutz, *et al.*, *Experimental Design: Paradigms and Procedures* (Inglewood, Calif.: Southwest Regional Laboratory for Educational Research and Development, 1967); and D. T. Campbell and J. C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," *Handbook of Research on Teaching*, ed. by N. L. Gage (Chicago: Rand McNally and Co., 1963), pp. 171-246.

necessary to determine the pre-experimental equivalence of the groups. The pretest means for the groups can be compared using appropriate techniques (t-test or analysis of variance) to establish that the groups are adequately similar to make valid generalizations based on the posttest data.

Another design that may be used is the Pretest-Posttest Control Group Design. The major difference between this design and the previous design is that individuals can be randomly assigned to groups when the Pretest-Posttest Control Group Design is employed. The design may be illustrated by the following:

$$\begin{array}{ll} R & M \longrightarrow T \longrightarrow M \\ R & M \longrightarrow M \end{array}$$

After the individuals are assigned to the experimental and control groups, the pretest is administered to provide baseline data. The experimental group then receives the treatment while the control group receives no special treatment. The posttest is administered to determine the effects of the treatment. The use of the pretest, posttest and a control group suggest the name for this design.

The third design discussed here is the Posttest-Only Control Group Design. This design differs from the Pretest-Posttest Control Group Design in that no pretest is administered. The schematic presentation of this design is given below:

$$\begin{array}{ll} R & T \longrightarrow M \\ R & M \end{array}$$

This design is used when individuals may be randomly assigned to groups and pretests are inconvenient, inappropriate, reactive, or unnecessary. No pretest is necessary if it can be assumed that the individuals do not possess the desired terminal behaviors prior to the treatment or if the only concern is to determine the relative effectiveness of two ways of accomplishing the same thing. The random assignment of individuals to groups assures a high degree of pre-treatment similarity between the groups.

The fourth, and final design to be discussed here is the Interrupted Time Series Design. This design may be illustrated as follows:

$$M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow T \rightarrow M_4 \rightarrow M_5 \rightarrow M_6$$

When the treatment requires many observations, or when many observations are being made routinely, the Interrupted Time Series Design can be employed. Typically, the design would be used when periodic measurements are made with a group of individuals. Measurements made prior to the introduction are used to establish a baseline. Measurements following the introduction of the treatment provide information on the treatment's effects. The multiple measurements following the treatment can provide data on short and long term effects which could not be detected through a single posttest only. This design is especially easy to use when regularly kept school records provide an adequate measurement for a treatment.

To summarize the discussion about research designs, the Non-equivalent Control Group Design is appropriate for situations where random assignment of individuals is not possible and the comparability of groups can be assumed or demonstrated.

The Pretest-Posttest Control Group Design may be utilized where random assignment is possible, change from pre-treatment to post-treatment measurement is of interest, and the pretest is not significantly reactive with the treatment.

The Posttest-Only Control Group Design is appropriate where randomization is possible and there is a likelihood that the pretest is reactive with the treatment. The design can be employed to compare alternative educational materials or processes.

The Interrupted Time Series Design may be employed where measurement data are collected at regular intervals over an extended period of time. It is particularly useful in situations when only one group is involved in a project.

The selection of appropriate analytic techniques must be based on the kind of sample (number of groups) and the kind of data which are generated by the measurement. The accompanying taxonomy (Figure 6, "Simplified Taxonomy of Statistical Methods for Experimental Research") provides a listing of techniques that are appropriate for various numbers of samples and kinds of data.

In most school research and evaluation, the t-test or the analysis of variance is used to analyze the data that are collected. The t-test can be employed with one group to assess the significance of change in pretest and posttest performance. It may also be used to compare two groups on change from a pretest to a posttest. The t-test can be used only when data are based on a scale with equal intervals (interval or ratio scale).

Although some modifications may be made in the formula for computing the t-test depending on the equality of sample sizes and variances, the basic formula for the t-test is:

$$t = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\left(\frac{\Sigma x_A^2 + \Sigma x_B^2}{N_A - 1 + N_B - 1}\right)\left(\frac{1}{N_A} + \frac{1}{N_B}\right)}}$$

Where $\overline{X}_A$ = mean for scores in Group A

$\overline{X}_B$ = mean for scores in Group B

Σx_A^2 = sum of squared deviations for Group A

Figure 6

SIMPLIFIED TAXONOMY OF STATISTICAL METHODS FOR DATA ANALYSIS IN EXPERIMENTAL RESEARCH

KIND OF DATA	KIND OF SAMPLE		
	One-sample	Two-sample	k-sample
Enumeration (Nominal)	Binomial test (Siegel, p. 36) Chi-square one-sample test (Siegel, p. 42)	Fisher's exact probability test (Siegel, p. 96); Chi-square test for two independent samples (Siegel, p. 104)	Chi-square test for k independent samples (Siegel, p. 175)
Ranked	Kolmogorov-Smirnov one-sample test (Siegel, p. 47); One-sample runs test (Siegel, p. 52)	Median test (Siegel, p. 111); Mann-Whitney U test (Siegel, p. 116); Kolmogorov-Smirnov two-sample test (Siegel, p. 127); Wald-Wolfowitz runs test (Siegel, p. 136); Moses test of extreme reactions (Siegel, p. 145)	Extension of the median test (Siegel, p. 179); Kruskal-Wallis one-way analysis of variance (Siegel, p. 184)
Metric (Interval or ratio)	t-test for one-sample (Walker and Lev, p. 228)	Random samples t-test (Popham, p. 144); One-way analysis of variance (Popham, p. [illegible]); Analysis of covariance (Popham, p. 233)	One-way analysis of variance (Popham, p. 179); Analysis of covariance (Popham, p. 233)

Note: The number following each source reference indicates the first page in the source where the statistical method is described.

References: W. James Popham, *Educational Statistics: Use and Interpretation* (New York: Harper and Row, 1967).
Sidney Siegel, *Nonparametric Statistics: For the Behavioral Sciences* (New York: McGraw-Hill Book Company, 1956).
Helen Walker and Joseph Lev, *Elementary Statistical Methods* (New York: Holt, Rinehart and Winston, 1958).

*Adapted from: Richard E. Schutz, Ellis B. Page, and Julian C. Stanley, "Curriculum Guide for a Course in Educational Media Research" (Madison, Wisconsin: University of Wisconsin, Department of Educational Psychology, October, 1962; report of project under USOE Title VII, Contract B-236), p. 12.

Σx^2_B = sum of squared deviations for Group B

N_A = number of individuals in Group A

N_B = number of individuals in Group B

The following example shows an application of the t-test:

Group A		Group B	
Score (X)	X^2	Score (X)	X^2
16	256	12	144
10	100	9	81
7	49	5	25
6	36	3	9
6	36	1	1

$\Sigma X_A = 45$ $\qquad$ $\Sigma X_B = 30$

$\Sigma X^2_A = 477$ $\qquad$ $\Sigma X^2_B = 260$

$N_A = 5$ $\qquad$ $N_B = 5$

$\bar{X}_A = \frac{45}{5} = 9$ $\qquad$ $\bar{X}_B = \frac{30}{5} = 6$

$$\Sigma x^2_A = 477 - \frac{(45)^2}{5} = 477 - 405 = 72$$

$$\Sigma x^2_B = 260 - \frac{(30)^2}{5} = 260 - 180 = 80$$

Using the t-test formula on the previous page:

$$t = \frac{9-6}{\sqrt{\left(\frac{72+80}{5\text{-}1+5\text{-}1}\right)\left(\frac{1}{5}+\frac{1}{5}\right)}}$$

$$t = \frac{3}{\sqrt{\left(\frac{152}{8}\cdot\frac{2}{5}\right)}} = \frac{3}{\sqrt{(19)\ (.4)}}$$

$$t = \frac{3}{\sqrt{7.5}} = \frac{3}{2.8} = 1.07$$

By refering to a table of the *t* distribution, it is possible to determine whether or not the computed *t* of 1.07 is sufficiently large to conclude that the two means (9 and 6) differ significantly. (This procedure is described in most textbooks on statistical methods.[4])

The analysis of variance is used when three or more groups of individuals are compared on a set of scores, such as a change in pretest and post-test performance. The analysis of variance can be used only when data are scored on at least an interval scale.

The following example, which is extracted from *Educational Statistics: Use and Interpretation*,[5] illustrates the application of the analysis of variance using the following formula:

$$F = \frac{s^2_{between}}{s^2_{within}}$$

	Group A		Group B		Group C	
	Score (X)	X^2	Score (X)	X^2	Score (X)	X^2
	7	49	4	16	2	4
	10	100	6	36	2	4
	10	100	7	49	3	9
	11	121	9	81	6	36
	12	144	9	81	7	49
Total	50	514	35	263	20	102
Symbol	ΣX_A	ΣX^2_A	ΣX_B	ΣX^2_B	ΣX_C	ΣX^2_C

$$\Sigma X_{Total} = 50 + 35 + 20 = 105$$

$$s^2_{within} = \frac{\Sigma x^2_A + \Sigma x^2_B + \Sigma x^2_C}{(N_A - 1) + (N_B - 1) + (N_C - 1)}$$

$$\Sigma x^2_A = 514 - \frac{50^2}{5} = 514 - 500 = 14$$

$$\Sigma x^2_B = 263 - \frac{35^2}{5} = 263 - 245 = 18$$

[4] For an example, see Popham, pp 143-63; and Allen L. Edwards, *Statistical Methods* (2nd. ed; New York: Holt, Rinehart and Winston, 1957), pp. 209-12.
[5] Popham, pp. 180-86.

$$\Sigma x^2_C = 102 - \frac{20^2}{5} = 102 - 80 = 22$$

$$s^2 \text{ within} = \frac{14 + 18 + 22}{(5\text{-}1) + (5\text{-}1) + (5\text{-}1)} = \frac{54}{12} = 4.5$$

$$s^2_{\text{between}} = \frac{\frac{(\Sigma X_A)^2}{N_A} + \frac{(\Sigma X_B)^2}{N_B} + \frac{(\Sigma X_C)^2}{N_C} - \frac{(\Sigma X_{\text{Total}})}{N_{\text{Total}}}}{\text{Number of Groups} - 1}$$

$$s^2_{\text{between}} = \frac{\frac{50}{5}^2 + \frac{35}{5}^2 + \frac{20}{5}^2 - \frac{105}{15}^2}{2} = \frac{500 + 245 + 80 - 735}{2}$$

$$s^2_{\text{between}} = \frac{90}{2} = 45$$

$$F = \frac{s^2 \text{ between}}{s^2 \text{ within}} = \frac{45}{4.5} = 10.0$$

The computed F ratio, derived through the analysis of variance, would be referenced in a table of F values to determine whether or not it is significantly large. If it is, this would provide a basis for concluding that one or more of the three treatments is significantly better than the remainder. The procedures for performing the computations and making the statistical test are described in books on statistical inference and research design.[*]

The preceding comments are intended primarily to increase your awareness of the need to review thoroughly the research-evaluation design and analytic procedures that apply to your program or project activities before the initiation of implementation. The design and analysis cannot be superimposed effectively after project activities have begun. Early attention to these needs can assist in providing a sound basis for making judgements about the success of your efforts.

[*] Edwards, pp. 257-73; Popham, pp. 179-88.

BIBLIOGRAPHY

Buros, Oscar Krisen, ed. *Sixth Mental Measurements Yearbook.* Highland Park, N. J.: Gryphon Press, 1965.

____________. *Tests in Print.* Highland Park, N. J.: Gryphon Press, 1961.

Cronbach, Lee J. *Essentials of Psychological Testing.* 2nd ed. New York: Harper and Row, 1960.

Ebel, Robert L. *Measuring Educational Achievement.* Englewood Cliffs, N. J.: Prentice-Hall, 1965.

Educational Testing Service. *Test Collection Bulletin.* Vol. 1, Nos. 1-6 (1967); and Vol. 2, Nos. 1-6 (1968). Published bimonthly at Princeton, N. J.: Educational Testing Service.

Edwards, Allen L. *Statistical Methods.* 2nd ed. New York: Holt, Rinehart, and Winston, 1967.

Gage, N. L., ed. *Handbook of Research on Teaching.* Chicago: Rand McNally and Company, 1963.

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

Oppenheim, A. N. *Questionnaire Design and Attitude Measurement.* New York: Basic Books, 1966.

Popham, W. James. *Educational Statistics: Use and Interpretation.* New York: Harper and Row, 1967.

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

Stanley, Julian C. *Measurement in Today's Schools.* 4th ed. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Thorndike, Robert L., and Hagen, Elizabeth. *Measurement and Evaluation in Psychology and Education.* 2nd ed. New York: John Wiley and Sons, 1961.

Travers, Robert M. W. *An Introduction to Educational Research.* 2nd ed. New York: Macmillan Company, 1964.

Walker, Helen M., and Lev, Joseph. *Elementary Statistical Methods.* Revised edition. New York: Holt, Rinehart and Winston, 1958.

____________. *Statistical Inference.* New York: Holt, Rinehart and Winston, 1953.

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

9
DATA COLLECTION TECHNIQUES

by
Noflet D. Williams

DATA COLLECTION TECHNIQUES

Noflet D. Williams, *Director*
Institutional Research Office
Tennessee Technological University

It is a pleasure to be with you to share some of my biases and ideas in the field of education. The title of this presentation might have been "Data Collection Techniques—Some Cautions and Comments not Generally Verbalized About the Importance of Precise Data Gathering."

Although data collection is important, it is neglected in the current literature. This assignment has been very frustrating because virtually every writer for research journals makes the assumption that the data will be or has been collected correctly. I would question whether this is a valid assumption for many of the research and evaluation projects in school systems throughout the country.

Before I start out with any additional ideas, let me confess to being a "generalist." I am trying to learn all there is to know about data analysis systems, sub-systems, and all the jargon used to describe the short range or long range planning process. Today, I would like to give you several basic guidelines to remember as you plan for data collection. These guidelines are: (1) use good common horse-sense when making decisions on data collecting: when you are faced with a problem, always look at it from a sensible standpoint and make reasonable decisions; (2) be alert during data gathering sessions and note any unusual circumstances; (3) be skeptical of any summary table or data analysis that looks unusual or out of the ordinary; and (4) build in data checks to assure accuracy at the time of test administration and data gathering sessions. If you are administering a test, it is no trouble, when a student completes the test and turns it in, to check to make sure the student has answered every question on the answer sheet, and, in situations where only one answer is appropriate, to verify that the student has not placed two or more answers for a single question on the sheet.

These checks are visual checks that an individual teacher can make without any real knowledge of research and evaluation. They will improve the data gathering process and apply to all four phases of evaluation as outlined in your Tennessee State Plan: Status Evaluation, Planning Evaluation, Operational Evaluation, and Final Evaluation.

I would like to digress for just a moment and compliment the State Department of Education, Memphis State University, and the consultants who assisted in developing the *Design for Tennessee Assessment and Evaluation of Title III, E.S.E.A.* The State Plan will provide a systematic plan of evaluation for personnel who are out on the firing line. This is a step in the right direction.

Let us assume for a moment that the Tennessee State Plan represents a management system. This management system is designed to coordinate, evaluate, and assess Title III projects in Tennessee. This is a tremendous undertaking and there will be many trials and tribulations as the system is refined. If you think of the State Plan as a system, then you will also need to think of all Title III projects as sub-systems.

You will need a small data base of basic information at the state level before you start your local assessments. Information that is nice to have is expensive. Many projects have bogged down simply because the project director included extra items of information and said, "We might need this item, we might want to do something with this a little later." With limited funds available to you, you had better prune items of information to a bare minimum in the beginning.

I assume from reading the State Plan that you will eventually have a computer at the state level that will analyze information for smaller school districts. Developing a source document is expensive in addition to the expense involved in collecting correct information, verifying the information, and running your data checks before the computer ever gets a chance to manipulate the data. In addition to the cost factor, we also find that the greater the data volume, with a limited staff, the more inaccuracies you can expect. In two seconds the computer can take an error and include it in many different tables. There is not a thing you can do about it once it gets into tables unless someone suspects that an inaccuracy exists. Sometimes you will spot it, sometimes you will print and then find an inaccuracy.

Data gathered by the State Department of Education, the Census Bureau, and other organizations should be used with caution. Each organization has its own set of definitions to meet its own purposes and needs. If you intend to use the data, let me encourage you to study and understand the definitions used in gathering the information. There is a difference between net enrollment and average daily attendance. The same thing is true for the other items of information that you may want to use.

Statewide base line data should be examined before you start developing your own plan of action. Every evaluation and research project should have a dictionary of terms for all data elements. The State Plan provides a set of definitions and I am sure additional definitions will be added. As I read the State Plan, I looked at your Survey of Critical Educational Needs in Tennessee. Item two on personnel asked you to indicate your greatest personnel need. I would suggest that your greatest need in personnel is going to be partially determined by your own personal philosophy of what a teacher should be and how much training this teacher should have. Your greatest need will be for qualified personnel if you are looking for persons with a master's degree. On the other hand, if 30 or 40 per cent of your teachers are without a bachelor's degree, you may be interested in securing teachers with a bachelor's degree. I think that, in addition to identifying and defining the terms related to the initial question, additional clarifications need to be provided.

The group that I met with this morning mentioned the fact that the Ohio Education Association has pushed through the Ohio State Legislature a bill making teacher aides a part of their over-all personnel system with a license that must be acquired before an aide can be employed to assist the classroom teacher and guidance couselor. Each state will need to define the role of the teacher aide in today's classroom.

The 1970 census provides an example of how to gather large volumes of data. You will be using this data as background information for future Title III projects. Almost all of you rely upon this data for household salary information and characteristics of the population. I want to review this with you for two reasons: (1) to show you one of the latest data gathering techniques that is in existence today, and (2) to take a bit of the halo off the information from the standpoint of it being the gospel.

One of the things that bothers me about any new development is the fact that today a group of people make a decision and take action. Tomorrow this new method of operation becomes sacred and no one dares to challenge how or why we operate in this way. So I think we need to realize that this data collection technique is not something to become sacred. It will require constant improvement.

The 1970 census will be taken in a unique fashion. Fifty to 65 per cent of the census questionnaires will be mailed to the individual household, will be filled out by the household heads in each of the urban areas, and then will be mailed back to the Census Bureau. The post offices have been involved in a rather laborious check and double check to make sure they have census questionnaires for every household in the urban cities. This part of the census can be classified as a mail survey. In addition to the mail survey, information on rural household units will be collected through a personal interview as a follow-up to the census questionnaire mailed to the household.

I would like to show you an example from the document. The sex of each member of the household will be recorded. The form contains a series of small circles that will be darkened to indicate the sex of each individual. We refer to this document as an input document and a source document. The census questionnaire is designed for electronic data processing. The information for the entire source document will be placed on microfilm which can be scanned by the FOSDIC process. The microfilm will put the answers in reverse form; dark spots will appear as white spots. The FOSDIC process will be used to scan these white spots on the microfilm and will transfer the information to a magnetic tape. The information on the magnetic tape will be computer edited.

I am interested in the procedure they use to edit the information. The computer examines the information as it was reported. If it receives information about a female as the head of the household, other bits of information will be checked against this to ensure that the information is consistent.

This data checking procedure also includes filling data gaps. If a bricklayer who has worked thirty-nine weeks does not report his salary, the computer checks his age, training, level of education, etc., searches for a person with a similar number of weeks of employment as a bricklayer and inserts that salary in the data gap. So the computer actually edits and updates the information without going back to the household for information.

The data element sex is information which cannot be divided into smaller components. In the years ahead, we will gather basic data elements which will be manipulated by the computer. Comparable data elements must be based on comparable standardized definitions. This is why we need a dictionary of data elements so that we can have a set of definitions that can be used throughout the state. Aggregate data then becomes the pooling of individual information on the sex of this individual and many other individuals.

A coder is used to code answers to open-ended questions included in the census. The coders receive special training to ensure that all coders will code similar answers in the same way. In many cases, we use test items that are open-ended and then ask someone to code the information for us without giving them the necessary training. If we are to be accurate, we must eliminate this method of operation. Also, most of us look at the census information and think in terms of this being the information on every household. In the 1970 census, certain questions will be answered by 100 per cent of the households, other questions by 20 per cent, and some questions by 5 per cent. In the situation where a 20 per cent sample is used, the data will be multiplied by five to provide an estimate for the total population.

I think we need to realize that Tennessee should develop a data system. One of the reasons I was encouraged when I read the Title III evaluation design was the fact that it did refer to the use of the computer which would allow you to develop a data system. If you think of data as material in a hopper and if you put corn and peas and a few other items in the hopper, your cornbread will taste rather strange. Good corn meal is required to produce good cornbread and good, reliable data is needed if accurate evaluation is to be performed. I can assure you that you can develop the best plan and select the best statistical techniques and I can tear your evaluation plan to shreds by using improper data-gathering procedures.

I can not overemphasize the importance of data-gathering to the success of the evaluation process. In many cases, we are making educational decisions based on data that is improperly gathered. I would like to think in terms of Tennessee developing a data bank based on the internal storage of information without it ever having to be removed from the computer except as a printed report. The information can be stored on drums, magnetic tapes and discs. If we do a good job of collecting our data, edit the input forms, process the data, examine the output information, make educational decisions based on accurate reports, update the data, and then store it so that we can use it again, we will have moved a long way toward developing the kind of data system that is needed.

If the State of Tennessee needs a data system, we should think in terms of each Title III project developing a data system. A data system is needed to determine the status of each group of children at any given point in time, record the educational experiences of students, and then measure the outcomes. This will probably involve an exchange of information between school districts and the State Department. Your urban areas will already have computer installations. The rural school districts, the smaller sparcely populated counties, will need to depend on the State Department of Education for computer facilities.

Let us look for just a moment at the working relationship which might be developed between the State Department of Education and each one of the individual projects. By utilizing a central computer and using optical scanning devices, a special staff of programmers at the state level can assist individual project staffs with project planning and processing data. Joint planning will aid in coordinating state-wide efforts and at the same time the central staff will be performing a service to individual school systems.

Education leaders should be willing to analyze the over-all achievement level of all of the students or a sampling of the students from the State of Tennessee. For too long, our students have been underachievers and we have been patting the public on the back and telling them how fine our school systems are. I think of the researcher who when asked, "How's your wife?" said, "Compared to what?" This is the way it is with student achievement in the State of Tennessee. How do our students achieve compared to what other group of students? On the average, students in Tennessee do not perform well on standardized achievement tests. We have a long way to go to provide the educational experiences our children need.

A state-wide testing program should be developed for selected grade levels. I think that it is feasible to test a stratified random sample of students in an attempt to identify specific areas where Tennessee students do not perform as well as we think they should.

Let us look at some of the types of instruments and techniques that we might use in collecting data. (We will not spend a great deal of time on this because you had a very fine presentation this morning on standardized tests of various types.) Student achievement can be measured by using standardized tests and teacher-made tests. Until the teachers of the State of Tennessee learn to construct more valid and more reliable tests, we will be forced to use standardized tests to evaluate the progress of students. I would like to see an increase in the number of projects using valid and reliable teacher-made tests. Their use will increase slowly as evaluators, educators, and classroom teachers learn to plan and work together.

Additional data collection techniques include questionnaires, case studies, grading scales, sociograms, attendance records, participation records, check lists, and dropout counts. We have a battle going on in America today between those individuals who feel that research should be on a very limited and sophisticated plane and those who believe that it should involve more practical and less well-defined investigations. Evaluation and data analysis simply provide you with information which you can use subjectively in making decisions. To use a rating scale, a sociogram, or some other data-gathering technique to secure information, would be an improvement over some of the guesses we have made in past years. As we use these various instruments to secure information, we should attempt to refine each instrument and improve our evaluation process. Our primary objective is to improve the quality of education by making correct educational decisions.

Special skills are required to collect data; therefore, persons responsible for gathering data should have appropriate training. I would like to see the State Department of Education, or some other group, take the lead in developing short filmstrips to instruct persons in the art of data collecting. They could be used to remind a teacher that he should not walk into a classroom and set the stage for testing by saying, "Well, here you are. I don't know why, but we are going to take a test today." On the other end of the continuum is the situation where the teacher walks in and literally begs and pleads students to do the best they can on a test. We should seek the assistance of behavioral scientists in planning testing procedures which will allow a child to perform to the best of his ability during a testing situation.

Some of you are familiar with the mark-sensing technique. This is an illustration of a simple mark-sense card. A mark-sense card will usually include space for the student name, student identification number, course identification, and answers to thirty or forty questions. The mark-sense card or mark-sense answer sheet has bubbles outlined on the form which must be blackened if a machine is to punch a card or read the results of the test directly into the computer. Students must receive correct instructions on using the cards and answer sheets if accurate data is to be collected. The test administrator should do a visual check as each answer sheet is handed in to ensure that the questions have been marked in the desired manner.

I will close by offering some general cautions related to evaluation and measurement. One of these cautions involves timing. Just before or after a big ball game, the county fair, or some other unusual event, is no time to give a test or use some other data gathering instrument. When you schedule a data-gathering session, attempt to get the student into a normal situation. A high noise level and distractions from moving objects or persons should be avoided. I think that it is very important that you keep noise to a minimum. This should include keeping the public address system off the air.

If we were to start working on an evaluation project, we should make sure that certain logical activities occur. We should train the persons responsible for collecting data to make sure that they understand collecting procedures. As soon as all of the collection instruments have been collected in one place, we should immediately have a visual edit performed to check for data gaps and incorrect information. The information should be key-punched and verified if we are not using source documents which can be scanned by machine. If we are using a source document that does lend itself to optical scanning, we should process the information, secure a printout, and do a visual check on the data. No assumptions about the accuracy of the information should be made until it has been checked in every possible way.

I am assuming that we have gone through the planning process which should have included several sessions with the computer programmer and evaluation specialist. At this point, we should review the source document and output tables to make sure that the programmed data checks have been included. Let me assure you that after we have completed all the visual and machine checks we might still have errors in the data. Output tables should be studied carefully in a final effort to spot inaccuracies. Errors can be kept to a minimum through careful planning, accurate data collecting, and proper statistical analysis.

10
WORKSHOP SUMMARY

by
Harry Bowman

WORKSHOP SUMMARY

Harry Bowman, *Research and Evaluation Assistant*
Southeast Educational Laboratory
Atlanta, Georgia

The final presentation at a workshop is generally a difficult assignment because the participants are usually more interested in ending the workshop than in hearing a speech. Consequently, my comments will be brief and will suggest some interrelationships among the topics presented by the workshop speakers.

It seems to me that we could review the several topics that have been covered at this three-day meeting and see the logical relationship among those topics, i.e., the ways that they compliment each other. I think that, if I had had the job of planning the order of the presentations, I would not change the sequence from what it has been.

First, we started out with an overview of evaluation and a recognition of the importance that is being attached to evaluation activities at this point in time. Then, a presentation was made concerning the Tennessee Evaluation Design, which in my opinion represents a very significant contribution to education in this state in that it takes some models which evaluation theoriticians have developed and applies those to a very practical evaluation problem—a state-wide assessment of education. I think there is much good information in the publication on the state-wide assessment that will indicate to you the areas of need that you might give attention to in terms of structuring projects that you may undertake.

Next, we heard a presentation on the topic of planning networks and program management. I think it is obvious to all of us that it is very important that an adequate amount of time and effort be spent on initial planning of projects and related activities, such as monitoring the implementation of those plans and periodic replanning as we find situations where things have not occurred as we have planned that they will occur. We find it necessary to apply this kind of principle in our operations at the Southeastern Educational Laboratory where we are developing new curricular materials, testing them with students, and revising them. A multiplicity of management factors are involved, such as gaining access to students when needed, being able to implement the evaluation plan, and allotting personnel time and resources to revise and retest materials. It is important to know in some detail something about all of the tasks that are involved in implementing any kind of a program or project so that effective implementation can be accomplished.

This presentation was followed by one dealing with behavioral objectives and the EPIC structure of variables. It is necessary for us to determine in advance of the implementation of a project specifically what we expect to be the results of our efforts. It should be emphasized that the structure of the evaluation is determined by the statement of objectives. The tasks and behavior that you want to change and the procedures that you are going to use to generate the change dictate, to a very large extent, the appropriate evaluation techniques and, subsequently, the appropriate techniques for analyzing the data that are collected.

The discussion of the behaviorial objectives was followed by a discussion of measurement instruments focusing on both available instruments which may or may not be standardized, and instruments that may be constructed by the classroom teacher or by the project staff. An argument could be made for a reliance upon standardized instruments since there is a scarcity of persons who have the technical knowledge to construct evaluation instruments. I think that we should concern ourselves with the business of developing this competency. It is usually possible to provide more valid instruments by having the project staff construct some measurement devices than by using available standardized instruments. I think it is preferable in many instances for us to construct our own measurement devices rather than using those which are already on the market.

The discussion of measurement instruments was followed by some comments on the application of appropriate analytic techniques. The types of data that are collected and the project design, for the most part, determine which techniques are appropriate. In terms of project design, I think that in most cases it would be preferable to establish an appropriate control group. In local school systems where a project is designed to include the entire population of students available, it may be possible to work out arrangements with neighboring school systems to provide an appropriate control group. This consideration should be reflected in the project design and the choice of analytic procedures.

The last presentation focused on data collection techniques and some of the problems that are associated with the collection of data. The quality of the data which are collected is critical from the standpoint of the validity of the analysis and the judgements that can be made about the effectiveness of any interventions that might be introduced in schools.

In conclusion, I think we can see that all of the topics that have been discussed at this workshop are components of the whole system of project design, implementation, management, replanning, and decision-making. In the final analysis, we want to be able to make some sound judgements about the effectiveness of the activities that are undertaken for the improvement of education.

11
GLOSSARY

by
Jack E. Miller

GLOSSARY

Jack E. Miller, *Assistant Professor*
College of Education
Memphis State University

This glossary is included to provide assistance in attaining a better understanding of the concepts contained in the papers presented. The words defined, exemplified, or characterized were selected to serve as both new and adjunct information. Examples are provided for additional clarity. Finally, certain definitions have contextual meaning, whereas others have a more universal meaning.

ANALYSIS
: Disassembly; a process of reducing by division, dissection, classification, partitioning, and similar actions the whole into its parts and subparts for better understanding. In other words, something may be analyzed into its parts. (Contrast with synthesis.)

ANALYSIS OF COVARIANCE
: A statistical method that extends the analysis of variance technique and adds features of regression analysis. This permits analysis of groups which are unequal on one or more independent variables. For example, if a lower ability group performed as well as a higher ability group on a measure, the technique of analysis of covariance would provide an analysis that would compensate for unequal ability. Thus, the effect of the treatment could be compared on the independent variable of ability.

ANALYSIS, OPERATIONS
: The reduction of an act or process of a particular endeavor, or series of acts, for a specified purpose into component parts for a better understanding of their operational relationships within the whole.

ANALYSIS OF VARIANCE
: A parametric statistical procedure that is used to analyze the variance that exists between groups (systematic variance) and within groups (error variance).

"BUZZ" WORDS
: Current set of terms (jargon) having reference to a specific topic and used by "experts" on the topic. For example, a person might refer to an expository textbook as built-in, orderly, organized knowledge.

CODING
: The process of changing data to a form that permits more efficient manipulation.

CONTROL
: A monitoring function designed to limit or restrain variations from a plan within the limits of predetermined criteria.

CONTROL GROUP
: A group that is the same as the experimental group except for the variable or variables being measured, which provides a basis for comparison.

In other words, the control group serves a monitoring function to make sure the plan is being effectively and efficiently carried out within stated criteria.

DATA
Plural of datum; a fact or statistic. For example, scores from a test, responses from a test, the number of pupils in average daily attendance, etc.

DATA, BASIC
Data reduced to one dimension and reported by the smallest unit. These provide a unit by unit description of status. For example, basic data on the public school teachers in Tennessee would consist of a description of each teacher along the variables identified (e.g., age, sex, years of teaching experience).

DATA CELLS
A mass storage, random-access data processing unit used with third generation computers.

DATA CHECK
Built-in, internal audits of data; the act or process of using prescribed standards or operations to verify, reject, or accept data. For example, when gathering data on public school teachers in a state, one would reject responses to items on questionnaires that extended above or below a logical or legal limit of age (e.g., below 18 and above 75).

DATA SYSTEM
Organized means and procedures for handling data.

EQUIVALENT GROUP
In research, a group is equivalent to another group if both groups were randomly selected from the same population. For example, in a study dealing with tenth grade students, the students placed in an experimental group are selected at random from all tenth grade students. The control group is equivalent because they, also, were selected at random from the same total population of tenth grade students.

EVALUATION
The process of determining relative value, importance, or effectiveness. Evaluation, as a process, extends from formative evaluation (What is happening or what has happened?) to summative evaluation (Why does it happen?).

EXPERIMENTAL GROUP
A group that is to receive the treatment. The control group and the experimental group are equivalent on all variables except the one to be measured. For example, if the effects of a "new" approach to teaching a course were to be tested and evaluated, the experimental group would receive the "treatment" of the "new" method, whereas the control group would be taught using the traditional method.

F DISTRIBUTION

A parametric description of a population that (1) is normally distributed, (2) uses interval measure, (3) has equal variance, (4) has means which are additive, and (5) is used only when observations are made independently. It is a theoretical tabular array of values arranged according to levels of significance and degrees of freedom. To be statistically significant, the computed F must equal or exceed the table value of F at the specified level of significance and the specified number of degrees of freedom.

FINITE

The property of having a definite beginning and end. (Contrast with infinite, which has no definite beginning or end.) For example, thirty students in a specific class at a specific time in a specific school is a finite population.

FLOW GRAPH MODEL

A graphic representation of the order in which functions or tasks have to be accomplished in order to reach the project objectives; a planning network with branches that connect at nodes or points. In such a network each branch originates at one node and terminates at the next node with direction from one node to the next being indicated by an arrow.

INPUT DOCUMENT

Any means used to transfer data from external sources into a data system.

INTERRUPTED TIME SERIES RESEARCH DESIGN

A research design in which ' .ated measurements are taken to establish the stability of measur past performances (baseline data). The treatment is administered, and an equal series of measurements are taken for comparison with the baseline data in order to establish the effects of the treatment.

Example: $O_1 O_2 O_3 X O_4 O_5 O_6$ X = treatment
O = observation or measurement

INTERVAL DATA

The second from the highest level of data (or measurement). The length or size of the interval is the same over the total distribution. A score of 75, for example, is as much greater than 74 as a score of 15 is over a score of 14.

KEYPUNCH

A device used to record onto paper cards or tape by punching holes in the cards or tape according to pre-determined format to represent alpha, numeric or special characters.

LEVEL OF SIGNIFICANCE

The probability that a difference of the derived (computed) magnitude could have occurred by chance alone; the size of the region of rejection of a hypothesis related to a distribution; an area (or level) designated as too far from the mean of the distribution for a computed value to fall and still be considered "like" those values within a specified area (level of confidence) around the mean of the distribution. For example, a .01 level of significance is another way of stating that the chances are 1 in 100 that the statistical difference found was due to chance alone. (Note: Statistical significance is not necessarily an important difference.)

MAGNETIC DISK

A random-access data storage device on which data are recorded on the surface of a magnetized, rotating disk. Data are recorded on and retrieved from the disks by electronic impulse through a "read-write" head.

MAGNETIC DRUM

A random-access data storage device that is a cylinder which accepts data recorded onto its magnetized surface.

MAGNETIC TAPE

A recording tape that has the capacity for containing data records which are recorded by electronic impulse and retrieved, processed or otherwise utilized or converted to data output by a computer.

MANAGEMENT

The act of making decisions or causing decisions to be made concerning the allocation of resources for the accomplishment of predetermined objectives; i.e., controlling personnel and other resources, planning, directing, organizing, budgeting, and reporting.

MANAGEMENT SYSTEM

A set of operating procedures used to acquire and to process data or information according to a predetermined rationale, including the presentation of the resultant information to decision-makers in a timely, meaningful form.

MARK-SENSE

The process or product of using a device to detect "marks" in special places and, then, to translate the marks into other input, output, or storage media.

MEAN (ARITHMETIC)

An average; a method of identifying the tendency of a distribution to centralize by the division of the number of factors (e.g., scores) into the added sum of all factors (scores) in the distribution. Example: Given three scores,

$$X_1 = 10, X_2 = 8, X_3 = 6,$$

by adding 10 + 8 + 6, we find a sum of 24. Next, the sum of 24 is divided by the three scores, resulting in a mean of 8.

MEASURE, INTELLIGENCE

A measure of the actual or potential acquired or native ability to deal with abstract concepts, learn from experience, and adapt to situations. In education, measures of scholastic aptitude are used as one means of determining the probability of an individual's chances of success in formal schooling.

MEASURE, INTEREST

A measure of the tendency to selectively attend (notice) a phenomenon. For example, an individual might be administered a vocational interest inventory to ascertain how closely that person's interest relates to the interest of individuals who are successfully employed in an occupation.

MEASURE, PERSONALITY

An assessment by inventory, projection technique, or simulation of the aggregate of human traits which determine each person's unique adjustment to his environment. Personality measures typically attempt to determine the nonintellectual aspects of an individual's psychological make-up.

MEASUREMENT

Act or process of determining quantity, dimensions, degree, etc., often in the form of a scale. Common forms of measurement in education are teacher-made tests or publisher produced tests.

MEASURES OF CENTRAL TENDENCY

Methods of identifying the point in a distribution which typifies the distribution; the point in a distribution along which a majority of the cases tend to fall or cluster. The mean, median, and mode are methods of identifying the points about which data in a distribution tend to centralize or cluster.

MEASURES OF VARIABILITY

Measures of the spread from, or the dispersion from, the measure of central tendency. Whereas the mean, for example, identifies the "pile-up" point in a distribution, the measure of dispersion identifies, by a single figure, the extent to which the scores are spread out from the measure of central tendency. Measures of dispersion are standard deviation, semi-interquartile range, and range.

MEDIAN (A PARTITION VALUE)

Midpoint; a measure of central tendency which identifies the point at which the number of factors (e.g., scores) above are exactly the same as the number of factors (e.g., scores) below; exactly separates the top half of the distribution from the bottom half.

Example: Given the scores 10, 8, and 3,

$$X_1 = 10, X_2 = 8, X_3 = 3,$$

the median is 8 because there is exactly one score above 8 and one score below 8.

MODE

A measure of central tendency which identifies the score that occurs most frequently in a distribution. Example: In a distribution of 10, 8, 6, 6, and 3, the score that occurs most frequently is the 6, which is the mode.

NETWORK SYSTEMS

Management tools which utilize the system's analytic and synthesizing procedures combined with the flow graph concept of nodes and branches (e.g., PERT/CPM).

NOMINAL (ENUMERATION) DATA

The lowest level upon which data may be classified (lowest level of measurement). Where data are used merely to classify or identify, those data are called nominal data (enumeration data) because a number used has no value as to its magnitude. For example, if a one is used for classification of males and a two is used for classification of females, it cannot be stated that a one is greater than a two. The numbers were used merely to classify or identify.

NONEQUIVALENT CONTROL GROUP DESIGN

A research design that uses a nonrandomized group as the control group. To have a randomized control group, the researcher must randomly select the subjects to be included in both the control group and the experimental group from a common population. The experimental group and the control group are not selected at random from a common population in the nonequivalent control group design. In education, groups (usually classes) are often taken as they are.

Example: $O_1 \; X \; O_2$ ···· = Lack of random assignment
··········· X = Treatment
$O_3 \quad O_4$ O = Observation or Measurement

NONPARAMETRIC STATISTICS

Statistical analyses which do not specify conditions about the parameters of the population from which a sample is presumed to have been taken. Nonparametric statistics are sometimes called "distribution free" statistics because they are not as limited to specified conditions of the distribution (or parameter values) as are parametric statistics.

OBSERVATION

An attempt to produce a synthetic description of reality. To observe with the intent of evaluating, the observer uses standards, has a purpose, and uses a predetermined procedure.

OBJECTIVE

Desired end considered worthy of attainment within the purposes for which activities are initiated.

OBJECTIVE, BEHAVIORAL

An objective stated in measurable terms describing the desired behavior which is intended to result from the activity. The objective is so stated that an unbiased determination can be made as to whether or not the desired outcome has been attained.

OPTICAL SCANNER

An input mechanism that translates and records data from marked documents by differentiating zoned areas that are marked from those that are not marked, or by otherwise translating characters or marks into a machine-acceptable form.

ORDINAL DATA

See Ranked (Ordinal) Data.

PARAMETRIC STATISTICS

Pertaining to parameter, which is an actual population value as contrasted with estimated values. When quantitative analyses are made using population measures, parametric techniques are used. When an analysis is made of samples of a population, statistical models are used.

Parametric statistics are based on assumptions about the population from which the sample has been drawn.

PLANNING
Outlining the future and/or deciding in advance what is to be done. The "purpose of a plan is to bring about behavior that leads to desired outcomes."

POPULATION
All of the members in a given group. The universe (population) may be either infinite—contains an unlimited number of possible members—or a population (universe) may be finite—containing a limited or specified membership.

POSTTEST-ONLY CONTROL GROUP DESIGN
A research design in which no pretest is administered. Both the experimental and the control groups are randomly equated. Measures other than a pretest can be used in this design.

Example: R X O_1 R = random assignment
R O_2 O = observation or measurement
X = treatment

PRETEST-POSTTEST CONTROL GROUP DESIGN
An experimental research design in which a pretest and posttest are administered to both the experimental group and the control group; the extent to which pretesting rather than the treatment affects the results can be measured.

Example: R O_1 X O_2 R = random selection
R O_3 O_4 O = observation or measurement
X = treatment

PRINTOUT
The printed output of data by a computer on paper in raw, synthesized, or analyzed form.

PROJECT
A specific task that is planned and undertaken to accomplish a specific result within specified time, budget, and performance specifications. Projects are finite, complex, homogeneous, and nonrepetitive.

PROJECT, CHARACTERISTICS OF ACTIVITIES
1. Finite in character. At a specific point in time, the objective or objectives will be attained and the project terminated.
2. Complex in nature. The combination and interrelationship of performance dimensions and tasks, schedules, resources, and budgets produce the complexities.
3. Unique set of acts which relate only to the project and make it possible to determine the line of demarcation between the project and its environment.
4. One time effort that terminates.

PROJECT MANAGER

Also known as project director or principal investigator. The person who performs managerial functions (see "Management") in coordinating efforts to the completion of a project (see "Project").

RANDOM ASSIGNMENT

One of the means used for controlling bias in research; to assign in a random or indiscriminate manner. Example: In the assignment of individuals to two or more groups, if each individual has an equal opportunity of being assigned to any one of the groups, the variation in the groups will not be due to the bias of the researcher because each individual will have been randomly assigned to each group on an "equal opportunity" basis.

RANGE

A measure of dispersion in which the highest score is reduced by the lowest score to yield the range or distance between upper and lower limits. (One variation is to subtract the lowest score from the highest score plus one, a process which is more pertinent to a distribution of ten or fewer scores than to a distribution of thirty or more scores.)

RANKED (ORDINAL) DATA

The next to the lowest level of data (or measurement). The numbers used indicate by their magnitude the order position of ranking. However, we cannot, using ranked data, state that a student ranked 2 is "as much better" than a student ranked 3 as a student ranked 5 is over a student ranked 6. Data that have interrelationships among all classes of >, i.e., greater than, higher than, more preferred.

Example: Sargeant > corporal > private.

RATIO DATA

The highest level of data (or measurement). This kind of data contains all of the advantages and properties of interval data, plus it has the property of "an absolute zero." A zero, using this level of measurement, means that there is absolutely none of the variable being measured present. For example, a zero on a properly calibrated and properly functioning instrument that measures weight would mean weightlessness. (Contrast this with a zero on a spelling test which would not mean a total absence of the ability to spell).

RELIABILITY, INSTRUMENT

An instrument that is valid and measures consistently is reliable. The concept of reliability centers upon the extent to which a valid instrument can be depended upon to obtain the same results, all other factors equal. The two major dimensions of the concept of reliability are time and instruments. See "Reliability, Types."

RELIABILITY, TYPES

The consistency with which an instrument measures may be determined by:

1. an internal comparison for consistency (one instrument at one setting).

2. correlating one form of an instrument with another form of that instrument.
3. correlating one instrument with an equivalent instrument.

The time lapse between administrations may vary from "immediately after" to "approximately a year," which plays a major role in determination of reliability.

SAMPLE
A part of the total population or universe.

SAMPLE, REPRESENTATIVE
A part of the total population or universe that is representative because each member of the population had an equal opportunity of being chosen.

SEMI-INTERQUARTILE RANGE
A measure of dispersion which is the quotient of the difference between the score representing the third quartile point and the first quartile point divided by two. The first quartile is the point at which the top three quarters of a distribution is divided from the bottom quarter of that distribution. The third quartile is the point at which the top quarter of the distribution is divided from the bottom three quarters of that distribution. The difference between the first and third quartiles divided by two represents the spread of the distribution around the second quartile (the median or fiftieth percentile).

SOURCE DOCUMENT
The instrument used to gather data (e.g., a completed questionnaire).

STANDARD DEVIATION
A single score which represents the spread, dispersion, or distance of each score from the mean of a distribution. To compute the standard deviation, subtract the mean from each score and square each. The sum of the squared deviations of each score from the mean is divided by the number of scores in the distribution (the number of scores minus one if there are fewer than thirty scores in the distribution); then, the square root of the result is the standard deviation.

Another approach to computing the standard deviation is to use the sums of squares method:

1. square each score
2. obtain the sum of each score squared
3. square the sum of the scores
4. divide the sum of scores squared by the number of scores
5. from the sum of each score squared subtract the sum of the scores squared and divided by the number of scores
6. divide the results of step 5 by the number of scores (the number of scores minus one if there are fewer than thirty scores in the distribution)
7. obtain the square root of the variance (the results obtained by performing steps 1-6).

STATISTICAL INFERENCE
The act of reaching conclusions about a population from one or more samples; to extrapolate from available data. For example, determining the probability that a sample mean came from a specified population.

SYNTHESIS
The utilization of the principle of assembly in such actions as summation, integration, unification, etc. The component parts are "put together" in a meaningful way for better understanding of the whole. For example, all the efforts in a school district to teach English might be better understood as a total effort rather than viewing each individual effort.

SYSTEM
The orderly (i.e., logical) arrangement of interdependent components or parts into a connected or interrelated whole to accomplish a specified goal. So defined, it is assumed that a system can be factored or resolved into a series of subsystems and each subsystem itself can be further factored or resolved. Typically, a system is a part of a larger system or has subsystems as components and has a specific task or tasks. For example, an instructional system in one school of a school district has its subsystems (e.g., curricula) and is also a part of the school district's total educational system.

SYSTEM ANALYSIS
Disassembly into parts, units, subsystems, activities, functions, and tasks of an objective-oriented whole into its component parts. For example, the instructional system of a school district might be analyzed subject by subject, educational level by educational level, school by school, and teacher by teacher. (Contrast with system synthesis.)

SYSTEM SYNTHESIS
System synthesis is that process of assembling an objective-oriented whole from its parts. In system synthesis, there is assembly into wholes, entities, networks, total systems, and flow diagrams. (Contrast with system analysis.) For example, the instructional system of a school district might be better understood as a whole, on some level, rather than as independent courses taught in different schools on different levels.

TECHNIQUE
A particular application of a method. For example, some techniques that might be used in applying the lecture method are to read from notes, not use notes, move about freely, remain in one place, provide duplicate outline of the lecture, etc.

t DISTRIBUTION (Student's t or Distribution)
A parametric description of normal populations which have the following characteristics:

1. normally distributed standard error (s) with mean zero
2. interval scale measure
3. equal variance or known ratio of variance

4. used only when observations are made independently (except for "correlated data").

It is represented by a theoretical tabular array of values that are divided into cells according to degrees of freedom and level of significance. A computed t is considered statistically significant if it equals or exceeds table value of t at a particular level of significance with a specified number of degrees of freedom.

TERMINAL BEHAVIOR

Behavior exhibited or available to be exhibited at the end of a treatment or activity. Simply put, the terminal behavior to be exhibited at the end of a shop teacher's unit of "properly driving a ten penny nail into hard wood" would be the act of a student performing as the behavioral objective prescribed: "properly driving a ten penny nail into hard wood."

t-TEST

A parametric statistic that is used to ascertain if the magnitude of the difference between two means is large enough to be considered statistically significant. Basically, the mean difference is divided by the square root of the group variance divided by the sample size.

TEST, ACHIEVEMENT

A measure of the actual rather than potential attainment of objectives. In contrast, aptitude tests measure the potential for achievement.

TEST, APTITUDE

A measure of the potential capacity or capability to acquire proficiency with education, training, and/or experience.

TEST, EDUCATIONAL

An examination of actual or potential cognitive ability which is evaluated by standard or standards of "correct or incorrect." As a means of measurement, a test is used to gather data for purposes of prognosis, diagnosis, or comparison. From the data gathered from a test, a teacher can make decisions concerning student performance and capability.

TESTS, EDUCATIONAL TYPE

The types of test used in education may be classified by:

1. producer: teacher-made or publisher-made
2. objectivity of scoring: objective (e.g., multiple choice) or subjective (e.g., essay)
3. level of cognitive processes required: memorization or reflective thinking
4. number of testees involved: individual or group
5. type of response: oral or written
6. level of difficulty: speed or power

In addition, the educational test may be classified according to how isomorphic the responses called for are, the degree of mastery required, etc.

TEST, STANDARDIZED

A measurement effort in which all variables are kept constant except the one being measured. A test is standardized as to the aspects of the test: construction, administration, scoring, reporting, and evaluation.

VALIDITY, CONCURRENT

The extent that an instrument agrees with other similar, currently available and accepted measures.

VALIDITY, CONSTRUCT

A measure of validity which treats the underlying psychological aspects of a test. The greater the degree to which the psychological aspects of a measure are in agreement with empirical data and are logically plausible the more valid is the instrument in construct.

VALIDITY, CONTENT (FACE)

The extent to which the content measured by an instrument is representative of the body of knowledge (or skills) that the instrument is purported to measure.

VALIDITY OF A MEASURE

The extent to which what an instrument measures or predicts can be known is the focus of the concept of validity. The greater the degree to which an instrument measures what it is supposed to measure, and nothing else, is the extent to which that instrument is valid. The types of measures of validity are content, concurrent, predictive, and construct (item structure may be included).

VARIANCE

The square of the standard deviation. The variance may be used in measuring the variation between groups (systematic variance) and within groups (error variance). The quotient of the sum of the squared deviations from the mean divided by the number of scores (the number of scores minus one if there are fewer than thirty scores in the distribution.) The variance may be determined by the sums of squares method. From the sum of each score squared is subtracted the quotient of the square of the sum of scores divided by the number of scores. The result is then divided by the number of scores (the number of scores minus one if there are fewer than thirty scores in the distribution).

$$\frac{\Sigma x^2}{n-1} = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}{n-1} \qquad V = \frac{\Sigma x^2}{n-1}$$

VERIFIER

A device used to "check" the transfer of data from source documents to input documents by regenerating the data record and comparing it with the input document.

12

PROGRAM OF THE EVALUATION RETREAT WORKSHOP

117

EVALUATION WORKSHOP RETREAT

July 28, 29, 30, 1969

Montgomery Bell State Park
Burns, Tennessee

Jointly sponsored by:

The Tennessee State Department of Education

and

Memphis State University

Workshop Director:

Fred K. Bellott

SUNDAY, JULY 27

4:00-9:00 P.M. Registration

6:00-8:00 P.M. Dinner

MONDAY, JULY 28

7:15-8:30 A.M. Breakfast

9:00-10:20 A.M. First General Session

Welcome: Devoy Ryan
Evaluation, An Overview: James W. Colmey
Tennessee Evaluation Design: Fred K. Bellott

10:20-10:40 A.M. Coffee Break

10:40-12:00 P.M.	First Small Group Meeting
12:00-1:00 P.M.	Lunch
1:15-2:15 P.M.	First Small Group Meeting (continued)
2:15-2:30 P.M.	Refreshments
2:30-4:30 P.M.	Second General Session

Planning Networks and Program Management:
Desmond Cook

4:30-5:30 P.M.	Individual Consultation
6:00-7:30 P.M.	Dinner
8:00-9:30 P.M.	Social Hour

TUESDAY, JULY 29

7:15-8:30 A.M.	Breakfast
9:00-9:40 A.M.	Second Small Group Meeting
9:40-10:20 A.M.	Third General Session

State Evaluation, A Team Effort: Roy Jones
Behavioral Objectives: Donald Thomsen

10:20-10:40 A.M.	Coffee Break
10:40-11:20 A.M.	Third General Session (continued)

EPIC Structure of Variables: Terry Cornell

11:20-12:00 P.M.	Third Small Group Meeting
12:00-1:00 P.M.	Lunch

1:15-2:30 P.M.	Third Small Group Meeting (continued)
2:30-2:45 P.M.	Refreshments
2:45-4:30 P.M.	Third Small Group Meeting (continued)
4:30-5:30 P.M.	Individual Consultation
6:00-7:30 P.M.	Dinner
8:00-9:30 P.M.	Social Hour

WEDNESDAY, JULY 30

7:15-8:30 A.M.	Breakfast
9:00-10:20 A.M.	Fourth General Session

Measurement Instruments: Willis Nowell
Application of Analytic Techniques: Harry Bowman

10:20-10:40 A.M.	Coffee Break
10:40-12:00 P.M.	Fourth Small Group Meeting
12:00-1:00 P.M.	Lunch
1:15-2:30 P.M.	Fourth General Session (continued)

Data Collection Techniques: Noflet Williams
Measurement Summary: Harry Bowman

2:30-2:45 P.M.	Refreshments
2:45-4:00 P.M.	Closing General Session

WORKSHOP PERSONNEL

Fred K. Bellott, Associate Director
Bureau of Educational Research and Services
Memphis State University

Harry Bowman, Research and Evaluation Assistant
Southeast Educational Laboratory
Atlanta, Georgia

Terry Bond, Project Evaluator
Memphis Community Learning Laboratory
Memphis City Schools

James W. Colmey, Director
Bureau of Educational Research and Services
Memphis State University

Desmond Cook, Director
Educational Program Management Center
Ohio State University

Terry Cornell, Evaluation Services Coordinator
EPIC Evaluation Center
Tucson, Arizona

Charles Epler, Administrative Assistant
Teacher Corps Project
Memphis State University

Linda Greer, Teacher
Memphis Community Learning Laboratory
Memphis City Schools

Tom Innes, Director of Research
State Testing Bureau
University of Tennessee, Knoxville

Roy Jones, Coordinator
Division of Instruction
Tennessee State Department of Education

Elma Mardis, Lead Teacher
Memphis Community Learning Laboratory
Memphis City Schools

Willis Nowell, Director
School Psychological Services
Tennessee State Department of Education

John Petry, Assistant Professor
Bureau of Educational Research and Services
Memphis State University

Devoy Ryan, Associate Dean
College of Education
Memphis State University

Donald Thomsen, Assistant Director
Central Midwestern Regional Educational Laboratory
Memphis Area Office

Noflet Williams, Director
Institutional Research Office
Tennessee Technological University

13
PICTORIAL PRESENTATION

Evaluation Workshop Retreat Speakers

Top: Harry Bowman
Bottom: Don Thomsen

Top: Noflet Williams
Middle: Willis Nowell
Bottom: Terry Cornell

Top: Roy Jones
Bottom: Desmond Cook

Small Group
Meetings

Photos by Ralph
Naylor, Tennessee
State Department
of Education