DOCUMENT RESUME

ED 071 342 EM 010 487

TITLE	Introduction to Psychology and Leadership. Research and Evaluation Plans; Part II.
INSTITUTION	Naval Academy, Annapolis, Md.; Westinghouse Learning Corp., Annapolis, Md.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	TR-6-3a
BUREAU NO	BR-8-0448
PUB DATE	9 May 69
CONTRACT	N00600-68-C-1525
NOTE	104p.; See also EM 010 418 and EM 010 419
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Autoinstructional Aids; Cost Effectiveness; Curriculum Design; *Curriculum Development; Data Processing; Individualized Curriculum; *Individualized Instruction; Individual Psychology; Instructional Design; Leadership; *Leadership Training; Management Education; Military Training; Models; Multimedia Instruction; Psychology; Research; *Research Design; Research Methodology; Social Psychology; Statistical Analysis; Technical Reports

ABSTRACT

In the second of a two-part document (part I is under EM 010 475), the Westinghouse Learning Corporation provides further details on the research to be carried out in the experimental leadership course instituted at the United States Naval Academy (the final report appears under EM 010 418, EM 010 419, and EM 010 484). Further details are presented on the experimental variables: presentation variables, media variables, student variables, and task variables. In addition, the plan of research and the experimental hypotheses are described. EM 010 418 through EM 010 447 and EM 010 451 through EM 010 512 are related documents. (RH)

Westinghouse Learning Corporation

RESEARCH AND EVALUATION PLANS - PART II

TR 6.3a

May [illegible], 1969

LEADERSHIP MANAGEMENT COURSE
RESEARCH AND EVALUATION PLANS - PART II

Contract No. N00600-68-C-1525

ABSTRACT

This report is the second of a two-part document which includes the research procedures and methodologies to be employed in evaluating the instructional system and learning modules of the Leadership Management Course instituted at the United States Naval Academy.

Approved by:

[signature]

Project Manager
Leadership Management Course

Westinghouse Learning Corporation
1840 Lomas Boulevard, Northeast
Albuquerque, New Mexico 87106

TABLE OF CONTENTS

TABLE OF FIGURES

Page

LIST OF TABLES

I. INTRODUCTION

This report discusses a number of studies which the Westinghouse Learning Corporation (WLC) is interested in conducting during the implementation of the experimental course at the United States Naval Academy (USNA). The results of these studies will not only be used directly in the continuing improvement of the Leadership Management Course, but will also have wide-range applications in the field of educational technology, specifically in the areas of materials preparation, media selection, program implementation, and classroom management.

The primary purpose of the proposed research is to evaluate the effects of different media and presentation forms on student performance. WLC plans to test a number of experimental hypotheses in which media and presentation design will be varied, while types of tasks and student characteristics are being identified. It will thus be possible to correlate student performance with both the characteristics of the learner and the characteristics of the task.

The assessment of student variables on a number of different scales (See TR-6.3a, Research and Evaluation Plans - Part I), the description of categories of learning tasks (Bloom, 1956; Gagne, 1965, Evans, Homme, and Glaser, 1962), and the definition of instructional variables along the dimensions of presentation (Tosti, 1968; Tosti and Ball, 1969) should provide sensitive and valid measures of the effects of these factors on the performance of students in the Leadership Management Course.

It is expected that experiments of this type, as part of a concentrated effort in educational research, will eventually result in a comprehensive understanding of the educational process, so that an instructor may choose with confidence the most effective media and presentation forms to teach a particular type of student a particular type of task.

2. EXPERIMENTAL VARIABLES

According to Tosti (1968):

> The quality of an educational system must primarily be defined in terms of change in student behavior... Every factor in the educational system must be evaluated as to its ability to modify, either directly or indirectly, the behavior of the student.

In designing such a behavior change system, several classes of variables must be considered. These categories are illustrated in Figure 1 below.

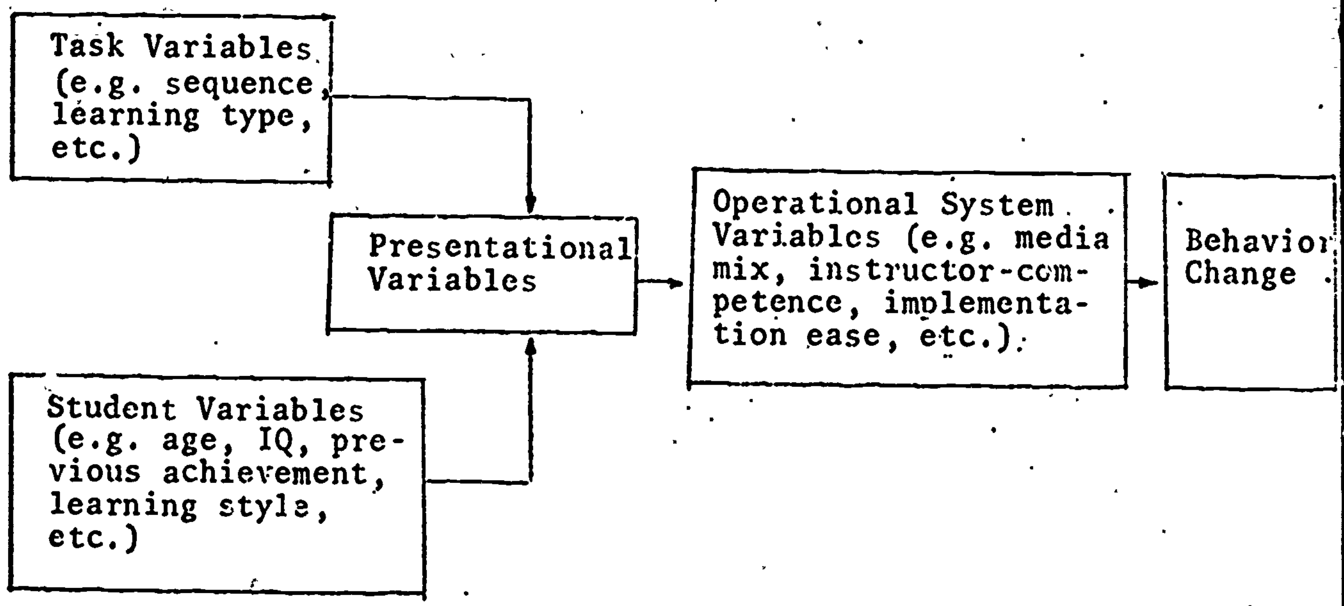

Figure 1. Categories of Instructional Variables

From Tosti's conceptualization, and from Figure 1 above, four categories of instructional variables, which are particularly important to the design of an instructional system, may be identified. They are:

a. Presentation variables

b. Media variables

c. Student variables

d. Task variables

In this section, each of these variables, and their relevance to the proposed research, will be described in detail.

2.1. PRESENTATION DESIGN (PRESENTATION VARIABLES)

One of the important questions in current educational research is, "Which media will teach a given unit of instructional material most effectively?" With the introduction of so many different technological aids (including teaching machines, programed texts, television, and film), there is today a steadily increasing variety of devices for use in any instructional program. The problem is to decide which medium is the "best" and what is the most efficient way to use it.

Studies which have attempted to analyze differences between media have been largely conflicting and ambiguous. Reid and MacLennan (1967), for example, reported 350 abstracts of media studies (mostly television and film); almost none of these studies found significant differences in media.

A number of researchers (Stolurow, 1962; Holland, 1965) have commented on the type of experimental comparison commonly attempted in studies related to programed instruction. Generally, the object of these experiments has been to determine the relative effectiveness of some existing instructional procedure compared to that of some new procedure or program.

This type of study has been criticized because the "existing instructional procedure, and often the new procedure or program, may be so ill-defined and poorly understood in terms of educational methodology that the results of any comparison are uninterpretable" (Ellis, 1962).

Ellis' criticism applies to most experimental designs in which different media are compared. Media may differ in a number of ways. A programed instruction (PI) text, for example, presents relatively small units of material at a time, requires active responding by the student, may provide immediate feedback (the correct answer), and allows for repetition of material. A film, on the other hand, is often viewed "passively" with a great deal of material presented at a time, and with no provision for feedback or repetition. Even if there are differences in the performance of subjects with these two media, it is impossible to specify which elements of the instruction are responsible. It appears that the central problem in a great many media studies has been the inaccurate identification of the instructional variables. This has resulted in "no significant differences" between treatment groups, as well as in hopeless experimental confounding.

In addressing this problem, Tosti and Ball (1968) have developed a model for presentation design and media selection in which a distinction is made between medium and presentation:

> Media researchers to date have not chosen to distinguish a presentation form from the media which carry it. The new model requires that such a separation be made.
>
> The media in instructional systems carry not only the data of the instructional message, but also data on students' responses and various bits of data necessary to maintain the operating systems. It is this conglomerate of information carried by a medium which will be called the presentation.

> Presentation forms will be explicitly structured to communicate all data (stimulus, response, system control, student control) necessary for an efficient student-system interface. A student does not learn from the media. He learns from the presentation form. Media do little more than deliver the information to be learned in whatever presentational form previously decided upon. Some media organizations have maintained that media choice may contribute to learning efficacy because of a student's media preference characteristics or because of media dependent cues. However, the importance of these two ideas is minimal when a separate presentation design is implemented.

To illustrate the distinction that Tosti and Ball have made, consider an instructor giving a lecture to one group of students. This lecture is videotaped and later shown to another group of students. Both groups would experience the same presentation design; only the medium would be different.

Also consider one lecture in which the instructor never answers questions, and another lecture where the instructor answers every question. In this case, the media are the same (lecture), but the presentations of the two lectures is different.

The media selection model which Tosti and Ball have developed is essentially a taxonomy of instructional presentation variables, independent of media device, content, and external constraints. Using this taxonomy, it becomes possible to precisely describe any instructional event by identifying its characteristics along basic dimensions which are common to all instructional presentations. Since the specification of instructional variables is a critical consideration in educational research, the application of Tosti and Ball's media selection model may result in a significant improvement in the quality of media studies and in the generalizability of their findings. The presentation taxonomy may be of comparable value to other educational classification systems. As Bloom (1956) commented on the taxonomy of education objectives:

> ...(the taxonomy) is expected to be of general help to all teachers, administrators, professional specialists, and research workers who deal with curricular and evaluation problems. It is especially intended to help them discuss these problems with greater precision.

The research planned for the Leadership Management Course may be considered a test of the utility of the media selection model. In each of the experimental hypotheses to be tested during the Leadership Management Course, the experimental treatments have been defined with reference to the six dimensions of presentation based on Tosti and Ball's media selection model.

2.1.2 The Dimension of Presentation

The six dimensions of presentation have been derived by a logical analysis of instructional systems. These systems possess three basic capabilities:

a. The transmission of instructional information (stimulus capability)

b. Accepting measurable behavior of the student (response capability)

c. Changing the presentation based on the behavior of the student (management capability)

These capabilities may be further defined in terms of two attributes: form and frequency. The result of this analysis is a 3 x 2 matrix from which the six dimensions of presentation have been generated. The matrix is illustrated in Figure 2 on the following page.

The six dimensions of presentation are:

a. Stimulus representation (stimulus form)

b. Duration (stimulus frequency)

c. Response demand (response form)

d. Response demand frequency (response frequency)

e. Management type (management form)

f. Management frequency

The following subsections give a detailed description of each of these dimensions.

2.1.2.1 <u>Stimulus Representation (stimulus form)</u>. This dimension is related most directly to the media form. It characterizes the dominant mode of sensory reception (by the student) of the instructional material. There are three categories within this dimension:

a. Verbal-written -- written material, such as printed text

b. Verbal-spoken -- voice transcriptions, such as from a lecturer, videotaped lecture, or tape recorder

c. Pictorial -- illustrative material, such as pictures

ATTRIBUTES OF SYSTEM CAPABILITY

SYSTEM CAPABILITY

	Form	Frequency
Stimulus	stimulus representation	duration
Response	response demand form	response demand frequency
Management	management form	management frequency
	Dimensions of Presentation	

Figure 2. The Dimensions of Presentation as Represented Within the System Capability Matrix

In many presentations, two or more stimulus forms may be used simultaneously. A book may display both illustrations and prose. An educational television program conveys both a picture and lecture. Other presentations can require media-mixes such as the teacher-blackboard combination. Less common is the simultaneous employment of two variations of the same stimulus presentation, i.e., requiring the student to read and listen to the same verbal presentation.

2.1.2.2 Duration (stimulus frequency). Tosti and Ball (1968) have explained stimulus frequency as follows:

> Presentation varies on this ordinal dimension from transient to persistent, depending upon the duration of the stimulus. Movies usually are conveyors of more transient presentation, and texts display relatively persistent ones. A classroom presentation by lecture is more transient than one which is delivered by the blackboard.

Transient presentations are usually instructor controlled. As in most films and lectures, the stimuli are available to the student for a fixed period of time. Persistent presentations are usually student controlled. An example would be the PI text, in which the student proceeds at his own rate and may study a unit of instructional material for as long as he likes. Essentially, a presentation form is categorized as "transient" or "persistent" depending on the length of time the presentation stimuli remain unchanged.

2.1.2.3 Response Demand (response form). This dimension characterizes the types of behaviors which students are expected to perform in an instructional situation. The four categories within this dimension are:

a. Covert
b. Overt-written
c. Overt-vocal
d. Passive

In a PI text, the student is asked to write the answers to small units of materials. This presentation design has an overt-written response demand. A student who is asked a question in a group discussion usually answers in the overt-vocal form. The covert category describes situations where the student is asked a question, but is not required to answer with a specific, overt (observable) response. For example, the instructor giving a lecture might say, "Think of what would happen if we mixed sodium and water." The passive category describes those presentations in which questions are not overtly asked, and the student is not expected to respond with specific overt or covert responses. Many lectures and most films are in this category. The student is only required to look and listen. However, the label "passive" should not imply that the student is doing nothing; he may be thinking intently, formulating questions about the material, or taking notes. This behavior, however, is controlled by the student, not by the instructional material. Where it is the intent of the instruction to evoke relatively specific behaviors in the learner, that presentation is categorized as either overt-written, overt-spoken, or covert.

2.1.2.4 Response-Demand Frequency (response frequency). This dimension describes how frequently the student is expected to respond (overtly or covertly) in a given period of instruction.

A PI text normally has a response demand after every frame. A lecture or film may be presented with no response demands in the entire session (or module). In any medium, questions or problems may be interposed at various intervals during the instructional sequence. This presentation design would have some intermediate response-demand frequency. This variable may provide a "better" index of "step size" than any other.

In a temporal sequence of instruction, there are three general dimension categories:

a. High response-demand frequency -- relatively frequent demand for a response in an instructional sequence, such as in programed texts where a response is required in each frame

b. Medium response-demand frequency -- relatively moderate frequency of demand for a response, such as questions which follow ten minutes of video-taped lecture

c. Low response-demand frequency -- low demand for a response, as when a "review" question follows a chapter of textual material

2.1.2.5 Management Type (management form). Instructional management (choice of management type) can be defined as those activities involved in the decision to assign a specific learning exercise to a given student, based on the assessment of some behavior of that student. One common example of instructional management occurs when the teacher, who discriminates that a student is having difficulty with learning a particular skill, makes the decision to assign special homework or decides to provide individual tutoring.

The general logic of this activity, i.e., assessing behavior, selecting presentation, and then having the student engaging in new activity, can be extended to provide the foundation for rules employed in most new individualized instructional systems and computer-managed classroom programs.

Management involves three activities:

a. Appraisal of data

b. Selection of some assignment as a result of some decision based on the data

c. Specification of the various actions that may be assigned

These three activities take place in every instructional system.

2.1.2.5.1 Types of Management. It is evident that the three elements of instructional management, i.e., repertoire assessment, selection decision, and new activity, can vary in their composition, depending on the purpose of management. Tosti (1968) has isolated six purposes that may be achieved. These are:

a. Need management -- to ensure students receive only those materials which they require to meet their objective.

b. Achievement management -- to ensure all students have mastered the objectives of the segment.

c. Prescription management -- to ensure a given student receives the most appropriate materials to meet the objectives in terms of his individual characteristics.

d. Motivation management -- to ensure continual student contact with the materials and to increase student learning rate.

e. Enrichment management -- to provide for additional information relevant to objectives, but not necessary for their achievement.

f. Systems management -- to provide logistic support relevant to materials availability, scheduling, administrative grading, and maintenance of the operating instructional system.

The present research is primarily concerned with three of these management purposes: achievement management, enrichment management, and systems management. The four catego.ies in the Management Type dimension are, therefore, the different procedures which may be used in management for achievement.

2.1.2.5.2 Management for Achievement. It frequently happens that a student is not responding to the presentation in a manner which allows him to reach the objectives. There are four strategic subclasses of management responses to such situations.

a. Repetition -- If the student fails to reach the objective, repeat the same presentation or continue through similar presentations until he does, e.g., individual prescribed instruction (IPI). Continuous practice is one variation of this strategy.

b. Multiform -- If the student fails to reach the objective with one presentation form, select a parallel but different form, e.g., Project PLAN (Flanagan, 1967).

c. Multilevel -- If the student fails to reach the objectives with the presentation form, select a lower level (more expanded) form, e.g., PROMOD (C'de Baca, 1968).

d. Error-Diagnostic -- If an error is made at any point within the presentation, action designed to correct that specific error is selected, e.g., intrinsic program presentation or computer assisted instruction (CAI) presentation. It is necessary when using the error-diagnostic strategy to classify errors as:

1) input errors -- due to poor presentation design.
2) processing errors -- due to the student's lack of the assumed appropriate repertoire on which the learning material was built, or the student's use of an inappropriate approach to the solution.
3) output errors -- due to carelessness, poor attention, and chance error (failed to attend to a significant stimulus).

The following is a summary typical of achievement management activities:

Appraisal	Selection Decision	Action
Examination of student error	Comparison of present level of achievement to final behavior objective	Student goes through alternate form of instructional task
Medium: curriculum imbedded tests	Medium: teacher or teacher aide	Medium: text or workbook

2.1.2.6 Management Frequency. "This dimension is ordinal and is ordered according to the relative frequency of the decision to modify the presentation" (Tosti and Ball, 1968). The concept of decision-making in presentation design is most clearly exemplified in tutoring. Typically, the student is directed to answer a question posed by the tutor, and a decision is made by the tutor about what he should next present, on the basis of that response. A similar instructional management form is used in PI. If an answer is correct, the student may be directed to ai one of a number of remediation frames.

Other media may also vary in decision frequency. An instructor may ask his class a question in the middle of his lecture to see if they are understanding the material. Depending upon the students' answers, the instructor may decide to continue with the planned lecture, to review the same material, or to start a new topic. For any presentation form, the decision frequency may vary from a decision every frame to no decision at all.

In a temporal sequence of instruction, there are three general categories:

a. High management frequency -- relatively high frequency of decision to alter the presentation, based on the student's response to a question. Management frequency may be built into the instructional system, as in a text where the decision is made on the basis of a response to every frame or to remediate him on the same frame. The management frequency may also be determined extemporaneously, as when a lecturer asks a class a question; if no one answers, the lecturer may decide to review previous content.

b. Medium management frequency -- relatively moderate frequency of decision to alter the presentation based on the student's response, such as having a quiz after a 10-minute film, and on the basis of the student's score, either repeating the film or proceeding to new material.

c. Low management frequency -- relatively low frequency of decision to alter instruction based on the student's response to a question, such as a lecturer giving a quiz after 40 minutes of lecture; basing the decision on the student's score, the instructor either assigns homework problems or does not.

It should be noted that the response-demand frequency must be equal to or more than the management frequency; one can't make decisions about a response more frequently than one calls for that response. An example of a presentation in which response-demand frequency exceeds management frequency is the lecturer who frequently asks the class "rhetorical questions"; the lecturer does not change his presentation on the basis of the student's (covert) responses, yet he does call for those responses. In this case the response-demand frequency would be high but the management frequency would be low. (See Figure 3, next page.)

	FORM	FREQUENCY
	Stimulus Representation	**Duration**
Stimulus	Verbal-written Verbal-spoken Pictorial	Transient ↑ Length of time the presentation remains intact ↓ Persistent
	Response Demand	**Response-Demand Frequency**
Response	Overt-written Overt-spoken Covert	Infrequent a. low or zero b. intermediate c. high ↓ Frequent
	Management Type	**Management Frequency**
Management	Repetition Multi-level Multi-form Error-diagnostic	Infrequent ↑ Frequency of decision to change presentation a. low or zero b. intermediate c. high ↓ Frequent

Figure 3. The Dimensions of Presentation

2.2 MEDIA (MEDIA VARIABLES)

As mentioned earlier, Tosti and Ball (1968) have theorized that student performance is determined largely by the presentation design, and that media alone exerts little effect on learning. For this reason, Tosti and Ball have not included media variables in their identification of the necessary and sufficient categories of instructional systems. (See Figure 1.) From this theory, however, a hypothesis about media effects may be derived: There should be no significant difference in student performance if media is varied while presentation design is held constant.

It is difficult to directly test this hypothesis, because positive results are based on "acceptance of the null hypothesis." It is statistically more desirable to obtain positive results based on "rejection of the null hypothesis." A large number of educational studies have found no differences in performance between various media conditions (Reid and MacLennan, 1967; Goldstein and Gotkin, 1962), and it is not clear what these results mean; because these experiments have confounded both presentations and media, they cannot be used to support Tosti and Ball's model.

In the present research, WLC intends to compare the (1) different media with the same presentation design, and (2) different presentation designs with the same medium. If significant differences are not found in the first condition, but are found in the second condition, the generalizability of Tosti and Ball's model will be greatly extended. The findings may also serve to redirect the general research effort in media; the question, "Which presentation design is more effective?" may be more important than the question, "Which medium is more effective?"

2.3 STUDENT CHARACTERISTICS (STUDENT VARIABLES)

A basic purpose of the proposed research is to associate the effects of instructional procedures to the characteristics of the learner. Three areas will be stressed in the evaluation of their relationship:

a. The isolation of student variables which are related to performance in specific media and presentation design forms

b. The isolation of student variables which predict academic success in the Leadership Management Course

c. The assessment of student preferences for specific media and presentation design forms

The rationale and suggested procedures for the research on student characteristics are comprehensively described in the first part of this document (TR-6.3a, Research and Evaluation Plans, Part I).

2.4 TYPES OF LEARNING (TASK VARIABLES)

It has been either specified or implied by a number of educational theorists that there are different types of learning which may be arranged in a conceptual order, from simple to complex.

Bloom, for example, has written:

> Although it is possible to conceive of these major classes (of behavioral objectives) in

> several different arrangements, the present one appears to us to represent something of the hierarchical order of the different classes of objectives. As we have defined them, the objectives in one class are likely to make use of and be built on the behaviors found in the preceding classes in this list.

Gagne (1965) has developed a taxonomy for classifying learning tasks into eight categories:

a. Type 1 -- signal learning

b. Type 2 -- stimulus - response connections

c. Type 3 -- motor chains

d. Type 4 -- verbal associations

e. Type 5 -- multiple discriminations

f. Type 6 -- concepts

g. Type 7 -- principles

h. Type 8 -- problem solving

These learning types can be structured in a hierarchy, so that if a given instructional sequence contains more than one type, mastery of the lower-order type is prerequisite to the acquisition of the higher-order type (i.e., problem solving (Type 8) requires as prerequisites, principles (Type 7), which requires as prerequisites, concepts (Type 6), etc.)

These task models are important, principally because they indicate that the effectiveness of an instructional presentation is partly dependent on the type of task to be learned, i.e., to produce optimal performance, different types of tasks may require different presentation forms.

Gagne (1965) has probably been the most specific in developing a set of prescriptions for the different kinds of tasks:

a. Stimulus-response learning: Provide repetition of the S $\longrightarrow$ R connection, applying reinforcement immediately.

b. Chaining (motor): Reinstate S $\longrightarrow$ R links in the proper order, either by beginning with the terminal link and working backward, or by using prompts. Reinforcement of the terminal link and also a certain amount of repetition are necessary.

c. Verbal chaining: Reinstate verbal links in the proper order, including mediating or "coding" links. Confirm correct responses. For longer chains, external "prompts" may be needed to provide sequence cues. Repetition is needed to overcome the effects of interference.

d. Multiple discrimination: Present stimuli in a manner that emphasizes distinctiveness. Repetition may be needed to reduce interference among individual connections. Confirm correct responses.

e. Concept learning: Present a suitable variety of stimuli to represent the concept class, each stimulus having a connection with a common response. Verify by presenting a novel stimulus that is also a member of the class.

f. Principle learning: Inform the learner of the performance to be expected. Invoke recall of components by verbal instructions. Make verbal statement of principle. Verify by the direction "Show me."

g. Problem solving: Inform the learner of the performance to be expected. Invoke recall of previously learned concepts or principles. Provide verbal guidance that directs thinking, but that does not state the principle being attained. Verify by asking the learner to "demonstrate" a concrete instance.

Gagne (1965) has summarized the rationale for coordinating presentation design to learning type:

> ...Decisions concerning what learning conditions are appropriate need to be applied to each learning act within a topic, and cannot be properly applied to a topic as a whole. Design of conditions for truly efficient learning can only be

achieved, therefore, by reference to a learning structure that identifies the individual learning acts of a topic and their relationships to each other

Since learning tasks may be stated as behavioral objectives (TP-6.2, Sequencing Rationale), WLC has developed a method for classifying the enabling and terminal objectives of the Leadership Management Course. This classification scheme, which has been derived and extended from Bloom's (1956) taxonomy, Gagne's (1965) learning types, and Evans, Homme, and Glaser's (1962) RULEG system, may facilitate the experimental evaluation of the relationships between learning type and presentation design.

Four types of behavioral objectives have been identified, and are listed below in order of increasing complexity:

a. Type 1 -- Definition - Identification

Given the instruction to identify the correct (purpose of definition of/description of/use of) concept X, the M will select from several choices the correct (purpose of/definition of/description of/use of) concept X.

(General Type 1)

1) Given the instruction to define concept X,
2) The M will select from several choices a definition of X
3) Similar to the following "X...."

(NOTE: The third part is optional.)

(Example Type 1)

1) When given the instruction to define "acquisition",
2) The M will select from several choices the correct definition.

(Variation:) a) Given several choices, the M will select the correct definition of acquisition.

(Variation:) b) (1) When asked to define "attention,"

(2) The M will select from several choices a statement

(3) (which indicates that attention is the selection of specific stimulus elements).

b. Type 2 -- Discrimination - Comparison

Given the instruction to evaluate the (relationship between/defining attributes of/contrast between/ comparison among) classes X, Y, Z...N, the M will be able to select from several choices the paragraph (which illustrates/describes/differentiates/identifies) this (relationship between defining attributes of contrast between/comparison among) classes X, Y, Z...N.

(General Type 2) 1) When asked to evaluate the relationship among X, Y,...N,

2) The M will select from several choices the paragraph which describes this relationship.

(Example Type 2) 1) When asked to evaluate the difference between retroactive and proactive inhibition,

2) The M will select from several choices, the paragraph which describes this difference.

c. Type 3 -- Generalization - Problem Identification

1) Deductive

Given examples of X, the M will be able to select from several choices the example which illustrates principle Y.

2) Inductive

Given an example of X, the M will be able to select from several choices which principle (X, Y,...or N) is (shown/exemplified/demonstrated) by the example.

(General Type 3)

1) Given examples of an X,

2) The M will be able to select from several choices the example which illustrates principle Y.

(Example Type 3)

1) When asked to compare several versions of the same communication,

2) The M will select from several choices the version which clearly links the subordinate's role to the overall objective of Naval operations.

d. Type 4 -- Problem Solving

When asked to evaluate a situation which is an example of class X, the M will select from several choices the correct (solution/approach/method/resolution of/reaction to) the situation using principle Y.

(General Type 4)

1) Given a problem situation which is an example of class X,

2) The M will select from several choices the correct approach to the situation,

3) Using principle Y.

(NOTE: The third part may be omitted if the objective is unambiguous. In practice, the third part is usually not given to the student.)

(Example of Type 4)

1) When asked to evaluate a situation in which there is an apparent failure in communication,

2) The M will select from several choices the description which indicates the appropriate action of a leader

3) Who assumes responsibility for the failure.

3. OVERVIEW OF THE RESEARCH PLANS

3.1 INTRODUCTION

This section is an introductory statement of the research planned for implementation in the Leadership Management Course. In the first subsection, the structure of the course is outlined, with particular attention given to a description of the cumulative post-test (CPT) Unit. In the next subsection, the first experimental hypothesis is described as a detailed example of WLC's design methodology. Following this subsection is a brief discussion of the methods for assessing the effects of the relevant experimental variables, and an overview of the proposed statistical procedures. The final subsection of this chapter includes general comments on the format of the research, some advantages of the proposed designs, and the possible constraints on research implementation.

3.2 GENERAL PLAN OF THE RESEARCH

3.2.1 Course Structure

As described in detail in the Course Strategy Report (TP-6.4), the Leadership Management Course has been organized in terms of the content and the temporal sequence of instruction. There are four basic elements in the course structure:

a. Part -- The content has been divided into 12 parts (or chapters). Each part is a formal designation of a relatively large topic area, representing a number of closely related terminal objectives. (See the Sequencing Rationale Report, TP-6.2.)

b. Segment -- Defined as content, a segment is a collection of learning objectives that are closely related by context and educational purpose. Defined temporally, a segment is a period of instruction, generally estimated to encompass 40 to 80 minutes of student learning time. Each part contains from two or ten segments, and there are approximately 60 core segments in the course.

   At the end of each segment, a progress check will be administered, to assess the students' achievement of that segment's objectives.

c. CPT Unit -- A CPT Unit is a group of three to five consecutive segments within a part. The CPT Units have been designated partly on the basis of content similarity across a number of segments, but primarily the designation is based on research requirements. The CPT Unit is the fundamental unit of instruction for testing the proposed experimental hypotheses. A cumulative post-test, administered after each CPT Unit to assess the achievement of that unit's objectives, serves as the most important dependent-variable measure of performance.

d. Module -- A module is an instructional presentation within a segment. Modules will serve as the basic instructional sequences in which different presentation designs will be directly compared.

As shown in Figure 4 below, Modules A_1, B_1, and C_1 are parallel modules within segment 1 (Modules A_2, B_2, and C_2 are parallel modules within segment 2, etc.). Within a given segment, the content is the same, but the parallel modules vary along particular dimensions of presentation, consonant with the experimental hypothesis being investigated.

CPT Unit

Seg. 1	Seg. 2	Seg. 3
Module A_1	Module A_2	Module A_3
Module B_1	Module B_2	Module B_3
Module C_1	Module C_2	Module C_3
PC	PC	PC

Pre-Test Over CPT — CPT

Figure 4. Structure of the Cumulative Post-Test (CPT) Unit

It should be noted that modules A_1, A_2, and A_3 ar- in the same presentation form (as are B_1, B_2, and B_3, etc.) The purpose of this design is explained in the following subsection.

3.2.2 Analysis of the Basic Experimental Design

To illustrate the basic experimental design using the CPT structure, consider the following hypothesis (Hypothesis I).

> In both simple and complex learning tasks, transient presentation with high response demand frequency (RDF) will be more effective than transient presentation with low RDF. There will be no difference in student performance between covert and overt response demand presentations, or between videotape and audiotape-panelbook media combinations

There are three sets of treatment factors evaluated by this hypothesis:

a. High and low RDF

b. Overt and covert response demand

c. Videotape and audiotape-panelbook media

Since the analysis of the RDF and response-demand (RD) dimensions are considered to be of primary importance, a suitable experimental design would be one in which each student would experience all of the possible treatment combinations. There are four such combinations:

a. High RDF and overt RD

b. High RDF and covert RD

c. Low RDF and overt RD

d. Low RFD and covert RD

These four treatments would be arranged as parallel modules within the segments comprising a set of four sequencial (but not necessarily consecutive) CPT Units. To test that part of the hypothesis dealing with the RDF and RD dimension, the students will be randomly assigned to one of four groups, and each group of students will receive each treatment condition in a different order. This design is technically a counterbalanced, repeated measures (within-subjects) variation of a "Latin-Square" (Lindquist, 1953). The design is illustrated in Figure 5 on the following page.

CPT UNITS

GROUPS OF STUDENTS (Rows)		Unit 1	Unit 2	Unit 3	Unit 4
	Row 1	high RDF, covert RD (module A)	low RDF, overt RD (module B)	low RDF, covert RD (module C)	high RDF, overt RD (module D)
	Row 2	low RDF, overt RD (module B	high RDF, covert RD (module A)	high RDF, overt RD (module D)	low RDF, covert RD (module C)
	Row 3	low RDF, covert RD (module C)	high RDF, overt RD (module D)	high RDF, covert RD (module A)	low RDF, overt RD (module B)
	Row 4	high RDF, overt RD (module D)	low RDF, covert RD (module C)	low RDF, overt RD (module B)	high RDF, covert RD (module A)

MODULE CHARACTERISTICS

Module A: high RDF, covert RD

Module B: low RDF, overt RD

Module C: low RDF, covert RD

Module D: high RDF, overt RD

Figure 5. Experimental Design for Evaluation of RDF and RD Effects

This basic design may be extended to permit comparisons between the third set of treatment factors (videotape and audiotape with panelbook). Two groups of students will receive all their presentations in the videotape medium; the other two groups of students will receive all of their presentation in the audiotape-panelbook medium. The final design, in which the main effects of the three sets of treatment factors and their interactions may be evaluated, is called a Repeated Measures Half-Plaid Square (Cochran and Cox, 1961, Plan 8.8, p. 332). This design is diagrammed in Figure 6 on the following page.

CPT UNITS

GROUPS OF STUDENTS (Rows)			Unit 1	Unit 2	Unit 3	Unit 4
	Videotape	Row 1	high RDF, covert RD (module A)	low RDF, overt RD (module B)	low RDF, covert RD (module C)	high RDF, overt RD (module D)
		Row 2	low RDF, overt RD (module B)	high RDF, covert RD (module A)	high RDF, overt RD (module D)	low RDF, covert RD (module C)
	Audiotape with Panel-book	Row 3	low RDF, covert RD (module C)	high RDF, overt RD (module D)	high RDF, covert RD (module A)	low RDF, overt RD (module B)
		Row 4	high RDF, overt RD (module D)	low RDF, covert RD (module C)	low RDF, overt RD (module B)	high RDF, covert RD (module A)

MODULE CHARACTERISTICS

Module A: high RDF, covert RD

Module B: low RDF, overt RD

Module C: low RDF, covert RD

Module D: high RDF, overt RD

Figure 6. Repeated Measures Half-Plaid Square Design for Testing Hypothesis I

3.2.2.1 Assessment of the Effects of Relevant Variables; Overview of Statistical Procedures. The CPT Test given at the completion of each CPT Unit, provides primary measures of student performance by which the effects of the relevant independent variables may be evaluated.

To assess the treatment effects and the interaction effects, an analysis of variance will be performed on the data. The analysis of variance source table for the statistical evaluation of Hypothesis I is illustrated in Figure 7 on the following page.

Source	IF Formula	IF	Error Term
Media	m-1	1	S/RM
Rows (R/M)	m (r-1)	2	S/RM
Subjects (S/RM)	mr (n-1)	44	———
Response Demand Frequency (F)	f-1	1	S X (FDU)/R
Response Demand (D)	d-1	1	S X (FDU)/R
F X D	(f-1) (d-1)	1	S X (FDU)/R
Media X F	(m-1) (f-1)	1	S X (FDU)/R
Media X D	(m-1) (d-1)	1	S X (FDU)/R
Residual	———	4	———
Units (FDM)	t-1	32	S X (FDU)/R
S X (FDU)/R	t (t-1) (n-1)	132	———

Figure 7. Analysis of Variance Source Table (simplified) for the Statistical Evaluation of Hypothesis I.

The relationships between student characteristics and performance on the various experimental conditions may be assessed by the correlational techniques which are described in TR-6.3a--Part I, and by the multiple-regression procedures described in the appendix of this report. The relationships between type of learning task and performance on the various experimental conditions may be assessed by classifying the data (performance scores on the CPT Tests) with respect to the type of objective. Since all four objective types will be represented in each CPT Test, four sets of scores may be abstracted, and a separate analysis of variance may be performed on each of them.

3.2.3 Concluding Statements and General Notes

The experimental design for testing Hypothesis I has been described in detail for the purpose of illustrating the basic elements of WLC's research methodology. Four of the five hypotheses planned for the Leadership Management Course will be evaluated by some variation of a counterbalanced, repeated measurements (within-subjects) Latin Square design. There are three important characteristics of this design which are worth repeating:

a. Each student experiences all the major experimental treatments.

b. Different groups of students receive the set of experimental treatments in different (counterbalanced) sequences.

c. The CPT Test, at the end of each CPT Unit, measures performance over the entire CPT Unit.

3.2.3.1 Advantages of Counterbalanced (Within-Subject) Designs

There are several major advantages of counterbalanced treatment arrangements which outweigh potential disadvantages in these experiments:

a. Within-subject evaluation of effects produces a marked increase in precision, compensating for limited sample size available, and tends to find treatment differences which are characteristic for individual students rather than for only some portion of the student sample.

b. Giving all treatments to each student roughly equates experimental histories, so that no individuals are disadvantaged in opportunity to obtain good grades in the course. This is a distinct administrative advantage, since experimental history need not be taken into account in assignment of grades.

c. Equation of experimental histories is advantageous in the assignment of students to further experiments, permitting randomization of students without treatment differences in prior experiments contributing to increased individual differences in following experiments.

d. The course content modules, with their CPT Tests, form natural blocks of experimental material which may differ widely in difficulty and performance level. Differences in modules are counterbalanced in these designs, while at the same time, investigation of treatment X module interactions is permitted. Thus, evidence of limitations on the generality of treatment effects is obtained.

3.2.3.2 Constraints on Research Implementation. In any non-laboratory setting, there are a number of problems which occur when implementing a research design. Discussed in this sub-section are difficulties which are both typical of classroom research and specific to research for the multi-media course development project.

Basic research in a classroom situation is not simple to accomplish. Having to fit the experiences into a planned, on-going course, multiplies the difficulties. The course requirements, fixed-class periods, limitations on testing, and lack of control between sessions all require some modification of the experimental design. The multi-media course development project has two primary goals: (1) to develop a successful Leadership Management Course, and (2) to examine basic variables which affect learning in general. These goals are not mutually exclusive. The results of the research on optimum presentation design variables will be used in the continuing improvement of the Leadership Management Course. The structure of the course sequence, however, will limit the manipulations of students and materials that are appropriate for a well-controlled evaluation of presentation variables.

Another research problem in any educational situation occurs because all of the students must learn a specific amount of course content. This limits the use of some control groups (students who do not receive any presentation) and the simultaneous investigation of more than two presentation variables (as "main effects") in a standard analysis of variance design. Other problems involve variation in achievement and base performance of students at different learning units.

Remediation procedures may be used to ensure that students will begin each new unit of material at the same level of achievement (performance), so that short-term effects are not confounded with effects of previous modules; there will still be some confounding when studying the long-term effects of different presentations.

The small number of students (50) participating in the course during the first year presents another difficulty for the planned research. A small subject sample severely reduces the sensitivity of the experimental comparisons (i.e., lowers the probabilities of finding genuine significant differences between treatment conditions), and greatly limits the feasibility of developing a large number of parallel modules. The paradox is that from other research and non-research considerations, 50 students is a very reasonable number with which to begin. There are bound to be a number of "bugs" in the first run of any large untested project, which will necessitate some changes in procedure and material. Because of this, it would be inadvisable to commit large numbers of subjects to the project the first time around. Yet, as the initial subject size is reduced for the sake of economy, so the probability of finding anything of value from the research is also reduced. This is a problem which confronts most researchers and innovators, and the best solution is the knowledge of its existence. The experiments planned for the first year at the USNA are not the end, but the beginning of the research effort; the results of the first year's study will not delineate final conclusions, but will serve as guides for future accomplishments. Viewed in this way, 50 students is a reasonable number with which to begin.

One final problem must be discussed. The students in the Leadership Management Course are all highly motivated, and the course itself will be constructed as an optimally effective instructional system. It may be that these two factors will result in relatively small differences between treatment groups. It may be difficult to obtain significant differences in comparisons between modules which vary in very subtle aspects, especially since most of the performance measures will reflect short-term learning. One solution to this problem will be to consider all the data (even if statistically non-significant) which indicate possible differences between treatment modules, as a basis for attempting a more complete and precise evaluation of the relevant hypotheses (with both a larger number of students and conditions). All non-empirical data (attitudes, preferences, comments by instructors and students, etc.) will also be analyzed.

From these considerations, it is apparent that research changes in the scope and logistics of the course will necessarily modify the types of research which can be conducted in the course. Consequently, a series of studies is being considered. Actual studies which can be conducted will be drawn from hypotheses when the constraints imposed by course materials are better known. The hypotheses to be presented will indicate relatively brief studies, each taking about two weeks to complete, and each concerned with analyzing small (but possibly significant) differences within one or two presentation dimensions. Later on, with more students participating, it will be possible to test more comprehensive hypotheses, involving the simultaneous variation of many values along many dimensions.

4. DESCRIPTION OF THE EXPERIMENTAL HYPOTHESES

4.1 INTRODUCTION

This section gives a detailed description of each of the five experimental hypotheses which WLC proposes to test in the first semester's implementation of the Leadership Management Course.

The hypotheses have been carefully constructed and revised many times, and were selected from a large pool of suggested studies. A number of important considerations were involved in their selection.

First, it was decided to conduct a number of studies, rather than only one or two, so that variations in presentation design and in other instructional procedures could be evaluated during the initial implementation of the course. Since one of the primary purposes of the research is to evaluate the utility of Tosti and Ball's (1968) presentation taxonomy, the first four hypotheses deal specifically with the expected relationships between student performance and presentation dimension variables. Hypothesis IV, however, is also concerned with evaluating the effectiveness of Csanyi's (1965,1961) "syndactic model." In Hypothesis V, WLC's elaboration of a "peer-monitor system" is analyzed. The five proposed studies will examine a variety of important research areas; it is expected that the experimental findings will provide direction for future research conducted at the USNA.

A second consideration concerned the construction of suitable designs for the sensitive evaluation of relatively few measures of performance with a comparatively limited sample size. (See

Constraints on the Research Design in the previous section.) These two factors indicated the use of within-subjects (repeated measures) designs to test four of the hypotheses. The remaining study (Hypothesis IV) is basically two between-groups experiments which may be combined to provide additional information.

A third consideration was the design of studies for which materials could be reasonably prepared and used. The final versions of the proposed experiments do not require the construction of excessively costly materials, nor the implementation of cumbersome procedures.

Finally, WLC has considered the problem of coordinating the research with the requirements of the course, so the students will profit from the experience. The use of counterbalanced designs, and the separation of treatment units, will help to ensure that the effects of the experimental conditions will be the same on all students. In addition, the comments and suggestions of students, instructors, and others associated with the course will be encouraged. The plan of instruction is sufficiently flexible to allow for appropriate revisions of materials and procedures.

A table listing the five experimental hypotheses is on the following page.

Table 1. THE EXPERIMENTAL HYPOTHESES

Hypothesis I

In both simple and complex tasks, transient presentations with high RDF will be more effective than transient presentation with low RDF. There will be no difference in student performance between covert or overt RD presentations, or between videotape and audiotape-panelbook media.

Hypothesis II

In both simple and complex learning tasks, persistent presentation with high RDF will be more effective than persistent presentation with low RDF. There will be no difference in student performance between three conditions of RD (overt-written, overt-spoken, and covert).

Hypothesis III

High RDF will be more effective than low RDF with either high, medium, or low management frequency.

Hypothesis IV

If performance is below criterion level with low RD presentation, remediation with high RDF will result in a significant improvement in performance.

Hypothesis V

Using "peer-monitor" procedures will significantly improve student performance in an instructional sequence.

4.2 THE EXPERIMENTAL HYPOTHESES

4.2.1 Hypothesis I:

> In both simple and complex learning tasks, transient presentations with high RDF will be more effective than transient presentations with low RDF. There will be no difference in student performance between covert or overt response demand presentations, or between videotape-panelbook and audiotape-panelbook media combinations.

4.2.1.1 Introduction. The primary purpose of Hypothesis I is to evaluate the effects of variations along two important dimensions of presentation (response demand frequency and response demand form) on student performance. This experiment will be carried out with two "transient" media: videotaped lecture and audiotaped lecture with panelbook.

In Hypothesis II, a "persistent" medium (text) will be used to test variations along the same dimensions of presentation, Since the goal of both hypotheses is to identify effective presentation conditions for the acquisition of knowledge, the introductory section to Hypothesis I is also relevant to Hypothesis II.

Tosti and Ball (1968) have stated a "rule" for determining whether a presentation should be transient or persistent.

> If one can invent a general rule, then it should be: The less the familiarity of the concept, the more persistent should be the presentation. With very difficult material, one should use a presentation that lasts as long as the student requires, i.e., a student-paced presentation.

Schramm (1964) has provided a brief review of some findings in this area. He has stated that although it makes sense intuitively that a student will learn more efficiently at his own pace, the experimental literature has not been able to demonstrate as much advantage for individual pacing as might be expected. He discusses the findings of several researchers in this area:

> Follettie (1961) found self-pacing better on an efficiency measure incorporating test score, training time, and testing time. Maccoby and Sheffield (1958) found that self-pacing worked best for superior students in learning from a training film interspersed with practice. On the other hand, no fewer than seven studies have found no significant difference attributable to individual or external pacing, whether the students are taught by teaching machines, programed texts, or television (Carpenter and Greenhill, 1963; Briggs, Plashinski, and Jones, 1955; Alter and Silverman, 1962; Feldhusen and Birt, 1962; Moore and Smith, 1961; Silverman and Alter, 1961; and Briggs, 1961).

One of the more interesting findings that Schramm discusses is that it seems to be possible to teach efficiently with programed materials on television or films:

> Carpenter and Greenhill (1963) compared an externally paced television program with self-paced teaching machine programs in three experiments, and externally paced films with a self-paced programed text in another. In each case they found no significant difference attributable to the pacing.

Even with variations in pace of 20 percent below and 10 percent above the average of class self-pacing, they found no decrease in the amount of learning.

Frye (1963) found evidence which not only furthered the understanding of pacing but indicated that the ability of the student may be a significant factor. He compared groups that were homogeneous in ability and found that a heterogeneous group took longer on the average to master a program when it was externally paced than when it was individually paced. In a homogeneous group, however, there was no difference between self-paced and group-paced instruction (Schramm, 1964). It may be that for the students at the USNA, the transience or persistence of an instructional sequence is less important than variations in other presentation dimensions, specifically RDF.

RDF was defined earlier as "how frequently the student is expected to respond (overtly or covertly) in a given period of instruction." It is the relative rate of response which is specifically elicited by the presentation design. In most programed instruction, the student is directed to respond in each frame (relatively high RDF). A lecture may proceed through an entire class period without requiring the students to answer a specific question (relatively low RDF).

It should be restated that the RDF is independent of the management frequency and management type dimensions. The answer to a question may or may not lead to a decision to change the presentation form, and an incorrect response may or may not be followed by remediation. An example of a presentation design with high RDF, but low decision frequency and no remediation, would be a film in which verbal-written questions are interposed at various intervals. Students would be directed to "think of the answer" to each question, and feedback (correct answer) might be given immediately. However, the film would continue regardless of what the students' answers were.

The present hypothesis is important for two practical reasons. First, it is difficult or impossible to have a high individual decision frequency with some forms of basically transient media (such as television, film, videotape, or audiotape). The use of these media could be greatly extended if it was shown that a high RDF by itself produces superior performance. Secondly, in most presentation designs a question is usually the occasion for a decision and for some kind of feedback. Yet feedback and remediation after the answer may be less important than the question. In a well constructed presentation design, the inclusion of appropriate questions at optimal intervals may be the most critical factor in producing superior performance. In fact, the Skinnerian model of linear programing is based partly on this principle.

A number of studies have demonstrated how presentation design may be improved by increasing the RDF. Studies by Gropper and Lumsdaine (1961) have shown that requiring active participation by students may produce superior performance in television presentations.

With particular types of subject matter, they found that the students who were presented programed instructional television scored significantly higher than students who were presented instructional television by conventional methods. They also found that significant correlations existed between the presentation form and the ability level of the students.

Part of their results have been summarized by Reid and MacLennan (1967):

> ...an ITV lesson on how movies work was shown in two versions. One group was a version which required active response and the lecturer supplied the correct answer after the student response.

> The other group saw the same lesson, but without active participation.
>
> Analysis of variance indicated that the active response students had significantly higher achievement scores on both the immediate and delayed post-tests than the passive group.

Abbey, et. al., (1963) also found that including participation questions in television instruction improved performance. They use three types of groups:

a. Groups who had overt responses to participation questions with feedback

b. Groups who had covert responses to participation questions without feedback

c. Groups who had covert responses to a conventional broadcast (having no participation questions) without feedback

Abbey found a significant interaction between overt and conventional responses, and related and correlated test items. The overt response group tended to do better on the related items and the conventional group tended to do better on the unrelated items. Subjects in the overt response condition tended to have higher achievement scores than those in the covert response condition to participation questions. The particular environment in which the students were tested (home vs. hospital and city) were also found to be important variables which were significantly related to response condition.

Greenhill (1967) has commented on the implication of these and other related studies. He has pointed out that there is presently a trend to incorporate some of the techniques of programed learning into television programs. This involves the inclusion of questions or short problems for students to solve, followed immediately by knowledge of results. Such an arrangement, he believes, provides for active audience participation, with immediate reinforcement.

He also implies that there will be an increasing amount of sophisticated research along these lines in the future.

It should also be noted that increasing RDF has the effect of increasing the number of "steps" in the program, and reducing the amount of material in each step. Schramm (1964) has described some of the problems in defining "size of step," and has discussed briefly the possible reasons for superior performance on small-step presentations (more practice and fewer errors). In general, Schramm (1964) states that "when significant differences have been found in learning from programs of different step size...they have usually been in favor of the programs with small steps." In the present study, the finding that high RDF is more effective than low RDF in transient presentation designs would support the reliability of Tosti and Ball's (1968) taxonomy, and would indicate the effectiveness of increasing the RDF in transient media.

Another experimental result which might significantly extend the utility of transient presentations is the finding of no difference in performance between overt and covert RD conditions. Although there were some differences in the Abbey, et. al., (1963) experiment, "the great majority of studies find no significant differences between the amount of learning from overt (constructed) and covert (non-constructed) responses" (Schramm, 1964). The particular value of the present research is that treatment effects can be assessed in "simple" and "complex" tasks; it may be that overt RD is more effective with complex tasks, and that neither RD type is more effective in simple tasks.

It should be noted that the two media to be compared in the experimental design of Hypothesis I are essentially in the same presentation form. Both videotaped lecture and audiotaped lecture with panelbook are transient media; in the present study both media will have zero management frequency, and therefore no management type. Both media are in the same stimulus form (verbal-spoken). Furthermore, in each individual CPT Unit, the audiotape-panelbook lecture will be the audio portion of the videotape lecture. The panelbook will be used to display such charts, diagrams, etc., as are indicated by the content material; in the videotape condition, these visual aids will be projected on the student's viewing screen.

As used in the present study, the two media are not completely "pure"; however, the visual aids will be only those essential to the instruction, and the results can be more reasonably generalized to contemporary instructional procedures, where mixing of stimulus forms is often desirable for practical reasons. As Tosti and Ball (1968) have written:

> ...In most presentations, two or more encoding forms may be used simultaneously. A book may display both illustrations and prose. An educational television program conveys both the picture and lecture. Other presentations can require media-mixes such as the teacher-blackboard combination.

4.2.1.2 Design and Procedure. The design for testing Hypothesis I has been described in Section 3.2.2, for the purpose of overviewing the basic rationale and methodology of the proposed research. In this section, that description will be amplified and completed.

For the sake of continuity, the outline of the design (Figure 6) is reproduced below.

CPT UNITS

GROUPS OF STUDENTS (Rows)		Unit 1	Unit 2	Unit 3	Unit 4
Videotape	Row 1	high RDF, covert RD (module A)	low RDF, overt RD (module B)	low RDF, covert RD (module C)	high RDF, overt RD (module D)
	Row 2	low RDF, overt RD (module B)	high RDF, covert RD (module A)	high RDF, overt RD (module D)	low RDF, covert RD (module C)
Audiotape with Panel-book	Row 3	low RDF, covert RD (module C)	high RDF, overt RD (module D)	high RDF, covert RD (module A)	low RDF, overt RD (module B)
	Row 4	high RDF, overt RD (module D)	low RDF, covert RD (module C)	low RDF, overt RD (module B)	high RDF, covert RD (module A)

MODULE CHARACTERISTICS

Module A: high RDF, covert RD

Module B: low RDF, overt RD

Module C: low RDF, covert RD

Module D: high RDF, overt RD

Figure 6. Repeated Measures Half-Plaid Square Design for Testing Hypothesis I

As shown in Figure 6, each student will be randomly assigned to one of four experimental groups. Two of these groups will receive videotaped lectures in each of the four experimental combinations of response-demand frequency and response demand. The other two groups will receive audiotaped lectures in each of the same presentation combinations.

In each CPT Unit, the audio portion of the videotaped lectures will be used as the audiotape presentation. Charts and other necessary visual aids will be represented on the viewing screen in the videotape condition, and will be referenced by the "off-screen" lecturer. In the audiotape condition, the lecture will be altered very slightly, so as to reference the visual aids in the panelbook.

In the high RDF modules, the lecturer will ask a question after each small "frame" unit of content (approximately twenty questions for each 50-minute segment). These will include simple questions on small units of content, and summary questions (criterion questions) on larger units of material.

In the low RDF modules, the lecturer will ask only the criterion questions during each segment.

In the overt RD modules, students will be instructed to write their answers on prepared answer sheets. In the covert RD modules, students will be instructed to "think of the answer," but not to indicate that answer in an overt form.

A summary illustration of the module specifications for Hypothesis I is shown in Figure 8 below:

MEDIA:	videotape or audiotape with panelbook (experimental variable)
PRESENTATION DESIGN	
Stimulus Representation	verbal-spoken
Duration	transient
Response Demand	covert or overt-written (experimental variable)
Response Demand Frequency	high or low (experimental variable)
Management Type	none
Management Frequency	zero

Figure 8. Module Characteristics for Experimental Design of Hypothesis I

4.2.1.3 Statistical Procedures. For the sake of continuity, the analysis of variance source table for the statistical evaluation of Hypothesis I (Figure 7) is reproduced below:

Source	IF Formula	IF	Error Term
Media	m-1	1	S/RM
Raws (R/M)	m (r-1)	2	S/RM
Subjects (S/RM)	mr (n-1)	44	——
Response Demand Frequency (F)	f-1	1	S X (FDU)/R
Response Demand (D)	d-1	1	S X (FDU)/R
F X D	(f-1) (d-1)	1	S X (FDU)/R
Media X F	(m-1) (f-1)	1	S X (FDU)/R
Media X D	(m-1) (d-1)	1	S X (FDU)/R
Residual	——	4	——
Units (FDM)	t-1	32	S X (FDU)/R
S X (FDU)/R	t (t-1) (n-1)	132	——

Figure 7. Analysis of Variance Source Table (simplified) for the Statistical Evaluation of Hypothesis I.

The following are two additional notes related to the design and statistical procedures:

a. The use of a quasi-Latin Square permits evaluation of the main effects of RDF, RD, RDF x RD, and the interaction of each of them with media. Information on the main effects of media is sacrificed by placement in rows.

b. While some information on the triple interaction is available in the residual component, the quasi-Latin Square, which sacrifices this interaction, permits the primary effects of interest to be investigated within an experiment of reasonable size.

4.2.2 Hypothesis II:

> In both simple and complex learning tasks, persistent presentations with high RDF will be more effective than transient presentations with low RDF. There will be no difference in student performance between three conditions of RD (overt-written, overt-spoken, and covert).

4.2.2.1 Introduction. Hypothesis II is the compliment of Hypothesis I in that it tests the effects of variations in the RDF and RD dimensions with persistent media (text). It should be recognized, however, that transient media will not be compared directly with persistent media for three main reasons. First, it is difficult to construct and to implement parallel modules which differ only in duration. Second, it appears more valuable to accurately assess the conditions for optimally effective presentations on each medium in separate experiments. Third, the use of two separate studies enables the effects of two other presentation dimensions to be evaluated in experiments of reasonable size.

In his discussion of the advantages and disadvantages of transient and persistent media, Tosti (1968) stated:

> The decision on the duration of the presentation is probably the least critical (but most often debated) one to be made in presentation design, since there are so many ways to compensate for it by manipulation of the other dimensions.

For this reason, the primary purpose of the Hypothesis II is to assess the effects of variations in the RDF and RD dimensions.

Three forms of RD will be compared in Hypothesis II, as a more comprehensive analysis of this presentation dimension. Although significant differences between RD conditions are not predicted, certain types of students may do better on one response form than on another. The repeated measures design constructed to test this hypothesis is particularly suitable for discovering these possible individual differences.

4.2.2.2 Design and Procedure. Two Latin Squares will be used to test Hypothesis II. This design is diagrammed in Figure 9 on the following page.

CPT UNITS

GROUPS OF STUDENTS (Rows)			Unit 1	Unit 2	Unit 3
	HIGH RDF	Row 1	overt-written RD (module A_1)	overt-spoken RD (module B_1)	covert RD (module C_1)
		Row 2	overt-spoken RD (module B_1)	covert RD (module C_1)	overt-written RD (module A_1)
		Row 3	covert RD (module C_1)	overt-written RD (module A_1)	overt-spoken RD (module B_1)
	LOW RDF	Row 4	overt-written RD (module A_2)	overt-spoken RD (module B_2)	covert RD (module C_2)
		Row 5	overt-spoken RD (module B_2)	covert RD (module C_2)	overt-written RD (module A_2)
		Row 6	covert RD (module C_2)	overt-written RD (module A_2)	overt-spoken RD (module B_2)

MODULE CHARACTERISTICS

Module A_1: high RDF, overt-written RD

Module B_1: high RDF, overt-spoken RD

Module C_1: high RDF, covert RD

Module A_2: low RDF, overt-written RD

Module B_2: low RDF, overt-spoken RD

Module C_2: low RDF, covert RD

Figure 9. Two Latin Squares Design for Testing Hypothesis II

As shown in Figure 9, each student will be randomly assigned to one of six groups. Three of these groups will each have only high RDF presentations; the other three groups will receive only low RDF presentations. The three groups of students experiencing each variation of RDF will be given the three RD treatments in a counterbalanced order. It can be seen that except for the RDF treatments, the set of rows 1, 2 and 3 are identical to the set of rows 4, 5 and 6, in terms of RD treatments. Each set of three rows is a Latin Square; hence the complete design is called Two Latin Squares.

The instructional media for all the modules will be a linear programed text with no confirmations. This medium commonly requires a student response in every "frame", which is specifically related to the content in that frame. Interpolated in the program are criterion questions, which assess acquisition over a number of frames.

The students in the high RDF condition will receive this basic linear program, with a response required in each frame. The students in the low RDF condition will only be directed to respond to the criterion questions.

In the overt-written RD modules, students will be requested to indicate their answers to these questions directly in the text-workbook. In the overt-spoken modules, students will be directed to verbalize their answers into a tape recorder. In the covert condition, students will be instructed to "think of their answers," and not to indicate them in any other form. It should be noted that within each CPT Unit, all the students receive exactly the same information, in basically the same program.

A summary illustration of the module specifications for Hypothesis II is shown in Figure 10.

MEDIA:	text
PRESENTATION DESIGN	
Stimulus Representation	verbal-written
Duration	transient
Response Demand	overt-written, overt-spoken or covert (experimental variable)
Response Demand Frequency	high or low (experimental variable)
Management Type	none
Management Frequency	zero

Figure 10. Module Characteristics for Experimental Design of Hypothesis II

4.2.2.3 Statistical Procedures. The analysis of variance source table for the statistical evaluation of Hypothesis II is shown in Figure 11, below.

Source	df Formula	df	Error Term
Frequency (F)	f-1	1	S/RF
Row (UD)	r-1	2	S/RF
Row X F (UDF)	(f-1) (r-1)	2	S/RF
Subjects (S/RF)	rf (n-1)	42	———
Units (U)	t-1	2	S X (UD)/RF
Demand (D)	d-1	2	S X (UD)/RF
U X F	(f-1) (t-1)	2	S X (UD)/RF
D X F	(f-1) (d-1)	2	S X (UD)/RF
Residual (UD)	r-1	2	S X (UD)/RF
Residual X F (UDF)	(f-1) (r-1)	2	S X (UD)/RF
S X (UD)/RF	rf (r-1) (n-1)	84	———
TOTAL		143	

Figure 11. Analysis of Variance Source Table (simplified) for the Statistical Evaluation of Hypothesis II

The following are notes relating to the design and statistical procedures:

a. Of primary interest, is the interaction of the RDF X RD treatments. The use of two Latin Squares permits the examination of the RD and RDF X RD interaction within subjects. The loss of information in the RDF is not serious.

b. Additional information is obtained on unit interaction with the RD and RDF X RD variables to permit examination of generality of effects. (If large interaction results, the study has less generality.)

4.2.3 Hypothesis III

> High RDF will be more effective than low RDF with either high, medium, or low management frequency.

4.2.3.1 Introduction. Schramm (1964) has pointed out that recent research has challenged many of the earlier assumptions underlying the construction of learning programs. According to Schramm, it has generally been assumed that an efficient linear program had:

> (1) an ordered sequence of items through which the student works in (2) short steps, therefore (3) making few errors as he records (4) a constructed response to each item, and (5) receives immediate knowledge of results.

However, Schramm cites three studies in which no difference was found in performance between logically sequenced and randomly sequenced programs. Smith and Moore (1961) found no differences in learning rate between spelling programs with different step sizes. Parry (1963) has argued that errors may facilitate learning in many situations and:

> the great majority of studies find no significant differences between the amount of learning from overt (constructed) and covert (non-constructed) responses.

In short, research has discovered many exceptions to the traditional "rules" by which programs have been constructed.

An important result of these experimental findings is that the rationale for developing and implementing various types of instructional systems is being seriously questioned. More specifically, the reasons for choosing one or another medium as an instructional device are being carefully reconsidered. As Tosti and Ball (1968) have written:

> Discussions of media usage for instruction have been favorite pastimes in education circles, and with the coming of popular writing on media by Marshall McLuhan and others, these discussions are now common in the popular press.
>
> Organizations devoted to the spread of new media information and excitement have appeared. The defense of the traditional media of lecture, laboratory, and text is also voiced. However, when their literature is examined from the standpoint of what is actually known about the learning process, unreliable data, faulty generalization from learning theory, and appeals to emotional or artistic bases are apparent.
>
> Proponents of broad and undefined media classes, such as audiovisual (AV), team-teaching, and computers (CAI, CMI), also seem to delight in throwing up some smoke to shroud their claims, and the other camps often respond emotionally. For example, if it were stated that most computerized instruction is a second-class imitation of tutoring or vice versa, various groups would immediately respond in anger. However, there is no current valid research that can support either claim.

In terms of the Tosti and Ball presentation taxonomy, the two management limitations have often been the basis of claims for the particular effectiveness of various types of "sophisticated" media (e.g., CAI). Proponents of these devices assert that they manage the behavior of the student more comprehensively, and that the increased "decision-making" capacity of the media increases the efficiency of the instructional system and results in superior performances by the students.

Yet, a number of studies cited earlier in this report do not support the assumption that a learner must experience a highly managed instructional presentation in order to achieve criterion learning. To paraphrase a previous statement, the frequency of decisions made after the student's answer may be less important than the question. In a well-instructed presentation design, the inclusion of appropriate questions at optimal intervals may be the most critical factor in producing superior performance.

As Tosti and Ball (1969) have written:

> The role of feedback in learning tasks is still a point of controversy among learning psychologists. However, there is general agreement that the law of contiguity is a primary force operating on learning in a given presentation, i.e., the student must make the response while attending to the relative elements of the stimulus for there to be a tendency for him to repeat that response when presented with that stimulus in the future. If the contiguous association is made, it is argued that the effect of positive confirmation will neither further increase nor decrease this tendency. However, the occurrence of negative confirmation (disconfirmation) evokes certain operants leading to the re-examination of the presentation (if it is still available), which may tend to suppress the just-emitted wrong response. In other words, the disconfirmation becomes the control stimulus for the subject to engage in some form of error self-management. The pos-

sibility of teaching students how to monitor their own behavior and how to correct their errors of carelessness is a conceivably more economical and universal solution than providing other decision media such as machine or teacher aids. In addition, one can conclude that it is not necessary to confirm every response.

In most experimental evaluations of learning programs and media systems, a question is usually the occasion for a decision and some kind of feedback. Thus, the independent effects of RDF and management frequency have not been adequately measured. In the present study, these measures can be made, using Tosti and Ball's taxonomy.

It should be noted that the primary purpose of Hypothesis III is not to compare different media. It is rather to assess the optimal conditions for learning within media. It is an underlying principle of Tosti and Ball's model, and of the present research, that the choice of media should proceed the discovery of optimal presentation design.

4.2.3.2 Design and Procedure. The design of Hypothesis III tests the effects of four variations of the RDF and management frequency (MF) dimensions in two different media forms. The two media are CAI and an audiotape-visual aid card combination (lectcard).

The CAI programs will be taught in the Computor Center at the USNA, where the student sits in a specially equipped carel, and receives instructional frames at a computor console which also is capable of accepting multiple-choice responses. In the lectcard medium, the student receives portions of a taped lecture through an earphone apparatus; he also is provided a visual aid card, which outlines the content of the lecture, and presents such charts, diagrams, etc., as are necessary.

There will be four treatment combinations of RDF and MF dimensions for each media unit. These are:

a. High RDF, high MF

b. High RDF, medium MF

c. High RDF, low MF

d. Low RDF, low MF

The outline for the between-subjects design to test Hypothesis III is diagrammed in Figure 12, below.

CPT UNITS (Media)

GROUPS OF STUDENTS		Unit 1 (CAI)	Unit 2 (Lect-Card)
	G_1	Module A	Module A
	G_2	Module B	Module B
	G_3	Module C	Module C
	G_4	Module D	Module D

MODULE CHARACTERISTICS

Module A: high RDF, high MF

Module B: high RDF, medium MF

Module C: high RDF, low MF

Module D: low RDF, low MF

Figure 12. Between-Subjects Design for Testing Hypothesis III

As shown in Figure 12, each student will be randomly assigned to one of four groups. Each group will be given the same presentation design in two (widely-separated) CPT Units. The presentation of each CPT Unit will be in a different medium (CAI and lectcard).

Both the CAI and the lectcard units will be intrinsic programs; the student will be given frames of material and questioned at some interval (depending on the RDF). If the student answers a question incorrectly, he may be branched (depending on the MF) to an error-diagnostic sequence in which the question is thoroughly reviewed and possible sources of error are analyzed. In the CAI unit, the frames will be illustrated on the computer console, and will be verbal-written material, augmented with necessary visual aids. The lectcard presentations will be a combination of verbal-written and verbal-spoken stimuli. The frames will be presented in short sequences of audiotaped lecture, with an accompanying outline (on cards) of the content (plus the necessary visual aids). The frame sequence for both media will contain two categories of questions:

a. Frame questions--covering the material in that frame
b. Criterion frame questions--which are interposed at intervals in the frame sequence, and which cover the content of a number of preceding frames

The basic characteristics of the four treatment modules for each media unit are as follows:

a. Module A (high RDF and high MF): The student is given a question in every frame (approximately 75 total frames per 50-minute segment, plus criterion questions). If the student answers any question incorrectly, he is immediately branched to an error-diagnostic remediation sequence which covers the same content as the incorrectly answered question.

b. Module B (high RDF and medium MF): The student is given a question in every frame, plus criterion questions. Regardless of whether the student answers the frame questions correctly, he proceeds to the next frame. However, if the student answers any criterion questions incorrectly, he is immediately branched to an error-diagnostic remediation sequence, which covers the same content as that of the incorrectly answered question.

c. Module C (high RDF and low MF): The student is given a question in every frame, plus criterion questions. Regardless of whether the student's answer to any question (frame or criterion) is correct or incorrect, he proceeds to the next frame. There is no remediation in this module.

d. Module D (low RDF and low MF): The student is given the frame sequence, but he is not asked any questions (either frame or criterion). This is essentially a conventional textbook presentation, in which the student is not required to respond to any specific question.

A summary illustration of the presentation characteristics for Hypothesis III is shown in Figure 13.

MEDIA	CAI and lect-card (experimental variable)
PRESENTATION DESIGN	
Stimulus Representation	verbal-written (CAI); verbal-written and verbal-spoken (lect-card)
Duration	persistent
Response Demand	overt-written (selected)
Response-Demand Frequency	high or low (experimental variable)
Management Type	error-diagnostic
Management Frequency	high, medium, or low (experimental variable)

Figure 13. Presentation Characteristics For Experimental Design of Hypothesis III

4.2.3.3 Statistical Procedures. Because the purpose of this experiment is not a comparison between CAI and lectcard, but rather an analysis of treatment effects within the two media, an independent evaluation of the treatment effects for each medium is planned. This may be accomplished by a simple one-way analysis of variance, followed by post-hoc comparisons between individual treatment groups (if significant F-values are obtained). Each media-CPT Unit is actually a separate study, in which the treatment effects are compared in separate between-subject designs.

Because the medium (and stimulus representation) is confounded with units, there is no interpretation possible for differences between CAI and lectcard. However, it is possible to get some additional information by analyzing the two studies together. The analysis of variance source table for this procedure is shown in Figure 14, below. Note that Treatment X Media Interaction can be evaluated if evidence from other experiments suggests the the Unit X Treatment Interaction is small.

Source	df Formula	df	Error Term
Treatments (T)	t-1	3	ľ
Subjects (S/T)	t (n-1)	44	--
Units (Media) (Units)	u-1	1	S X U/T
Units X Treatments	(u-1) (t-1)	3	S X U/T
S X U/T	t (n-1) (u-1)	44	

Figure 14. Analysis of Variance Source Table (simplified) for the Statistical Evaluation of Hypothesis III

4.2.4 Hypothesis IV

If performance is below criterion level with low RDF presentation, remediation with high RDF will result in a significant improvement in performance.

4.2.4.1 Introduction. One major goal of the preceding three hypotheses was to demonstrate that students will generally perform better on presentations with high RDF. The numerous exceptions to this general statement, however, have been the primary source of many criticisms recently levelled at linear programed instruction, which has high RDF as an elemental, identifying characteristic.

Although PI has been "the only media group which has made some effort to back up its claims with sound data" (Tosti and Ball, 1968), and although a number of studies cited earlier in this report indicate that "small-step" programs are generally more effective than low RDF presentations, there has been an embarrassing number of cases in which students do not like, and do not perform well on, linear programed instruction.

One of the chief compliants has been that some students are able to comprehend relatively large units of content material, and are thus bored by frame sequences in which responses to small pieces of information are required at frequent intervals. If it is true that individual students vary in their ability and preference for differing RED presentations, then this criticism of PI--that it fails to consider individual differences along the RDF dimension--is well taken.

Csanyi (1965, 1961) has suggested a procedure by which this apparent RDF limitation might be effectively dealt with. He has extended the format of the basic linear program so that students who do not require a frame sequence of instruction do not have to go through it. In this "syndactic-text" program, the student first receives, and is tested on, a relatively large unit of information presented in textual form. If the student passes the test on the summary statement, he proceeds directly to the next summary statement. However, if the student does not pass this test, he is immediately branched to a frame sequence covering the same material. Csanyi has reported that this technique produces superior performance in many different types of students.

Hypothesis IV is a formal test of the general principles underlying the effectiveness of the syndactic procedure: Students who are unable to perform well on low RDF presentations will significantly improve performance, if they are remediated on high RDF presentations. If results support this hypothesis, the generalizability of Tosti and Ball (1968) model will be increased, and the utility of Csanyi's syndactic format will be extended.

4.2.4.2 Design and Procedure. The design for testing Hypothesis IV follows the syndactic procedure; each experimental module consists of two basic parts: the summary statement and the remediation sequence.

Three CPT Units will be used, and each unit will present the summary statements in a different medium (text, slide-tape, and videotape). In each module, the student will receive summary statements in one of the above three media. If the student passes a test on the summary statement, he will proceed to the next statement. However, if he does not pass the test, he will be branched immediatly to one of three remediation procedures in the text medium. These different forms of remediation are the experimental treatments:

a. Module A -- text remediation in high RDF (in frame sequence, with a question in each frame)

b. Module B -- text remediation in low RDF (in rewritten summary statement form, with a test after the remediation statement)

c. Module C -- no remediation (control), the student proceeds to the next summary statement regardless of his answers on the proceding summary statement

As in the first two hypotheses, each student will receive each experimental treatment in each summary statement medium. Groups of students will be given the treatments in counter-balanced order.

The design for testing Hypothesis IV is diagrammed in Figure 15, on the following page.

CPT UNITS

	Unit 1 (summary statements in text medium)	Unit 2 (summary statements in slide-tape medium)	Unit 3 (summary statements in videotape medium
R_1	text remediation with high RDF (module A)	text remediation with low RDF (module B)	no remediation (module C)
R_2	text remediation with low RDF (module B)	no remediation (module C)	text remediation eith high RDF (module A)
R_3	no remediation (module C)	text remediation with high RDF (module A)	text remediation with low RDF (module B)
R_4	text remediation with high RDF (module a)	no remediation (module c)	text remediation with low RDF (module b)
R_5	text remediation with low RDF (module B)	text remediation with high RDF (module A)	no remediation (module C)
R_6	no remediation (module C)	text remediation with low RDF. (module B)	text remediation with high RDF (module A)

MODULE CHARACTERISTICS

Module A: text remediation with high RDF

Module B: text remediation with low RDF

Module C: no remediation (control)

Figure 15. Two Orthogonal Squares Design for Testing Hypothesis IV.

As shown in Figure 15, each student will be randomly assigned to one of six groups. In each CPT Unit, the students will receive the initial summary statements in a different medium. If a student does not pass a test following a particular summary statement, he will be branched immediately to one of three remediation conditions (including the "no remediation" control condition) in the text medium.

As an example of the experimental procedure, consider the order of events in Unit 1.

In Unit 1, the students in four of the groups (rows) will each be provided with two text workbooks, one containing the summary statement and the other containing the remediation sequences.

The students in rows 1 and 4 (Module A) will read a summary statement in the first text and then take a short test covering the material in that statement. If the student passes this test he will go on to the next summary statement. If he fails the test, he will be branched to the second text, which contains a rewritten version of the summary statement in a frame sequence with a response required in each frame. At the completion of this sequence, he will be given the same test and regardless of whether he passes it or not the second time, he will go to the next summary statement in the first text workbook.

The students in rows 2 and 5 (Module B) will receive the same summary statements and tests as rows 1 and 4. However, if the student fails to pass a summary statement test, he will be branched to the second text, which contains a rewritten version of the summary statement, which has all the mediators and other prompts as would a frame sequence, but a zero RDF (no questions). After reading this rewritten statement, the student will take the test again; regardless of whether he passes or fails this test the second time, he will go to the next summary statement in the first test. The students in rows 3 and 6 (Module C) will only receive the first texts with the summary statement. After reading a statement, a student in this condition will take the test covering the content of that statement. Regardless of whether he passes or fails the test he will be instructed to continue on the next summary statement.

Now consider Unit 2. The summary statements in this unit will be presented on a slide-tape apparatus, and each student will be given question books to answer the test which follows each statement. If a student in Unit 2, rows 3 and 5 (Module A), fails a summary statement test, he will be branched to a text containing a written version, in frame sequence, of the slide-test statement. If a student in Unit 2, rows 6 or 1 (Module B), fails a summary statement test, he will be branched to a text containing a rewritten version of the same statement, with all the mediators and other prompts included (but with no questions). The students in Unit 2, rows 2 and 4, will each receive the slide-tape presentation of the summary statements, and will take a test following each statement; however, regardless of whether they pass or fail any test, they will continue on the next slide-tape summary statement.

This procedure will be essentially the same in Unit 3, except the summary statement will be presented by videotape medium;

the remediation will again be rewritten versions of the summary statements, in the text medium.

An illustration of the procedure for Unit 2 is shown in Figure 16 below.

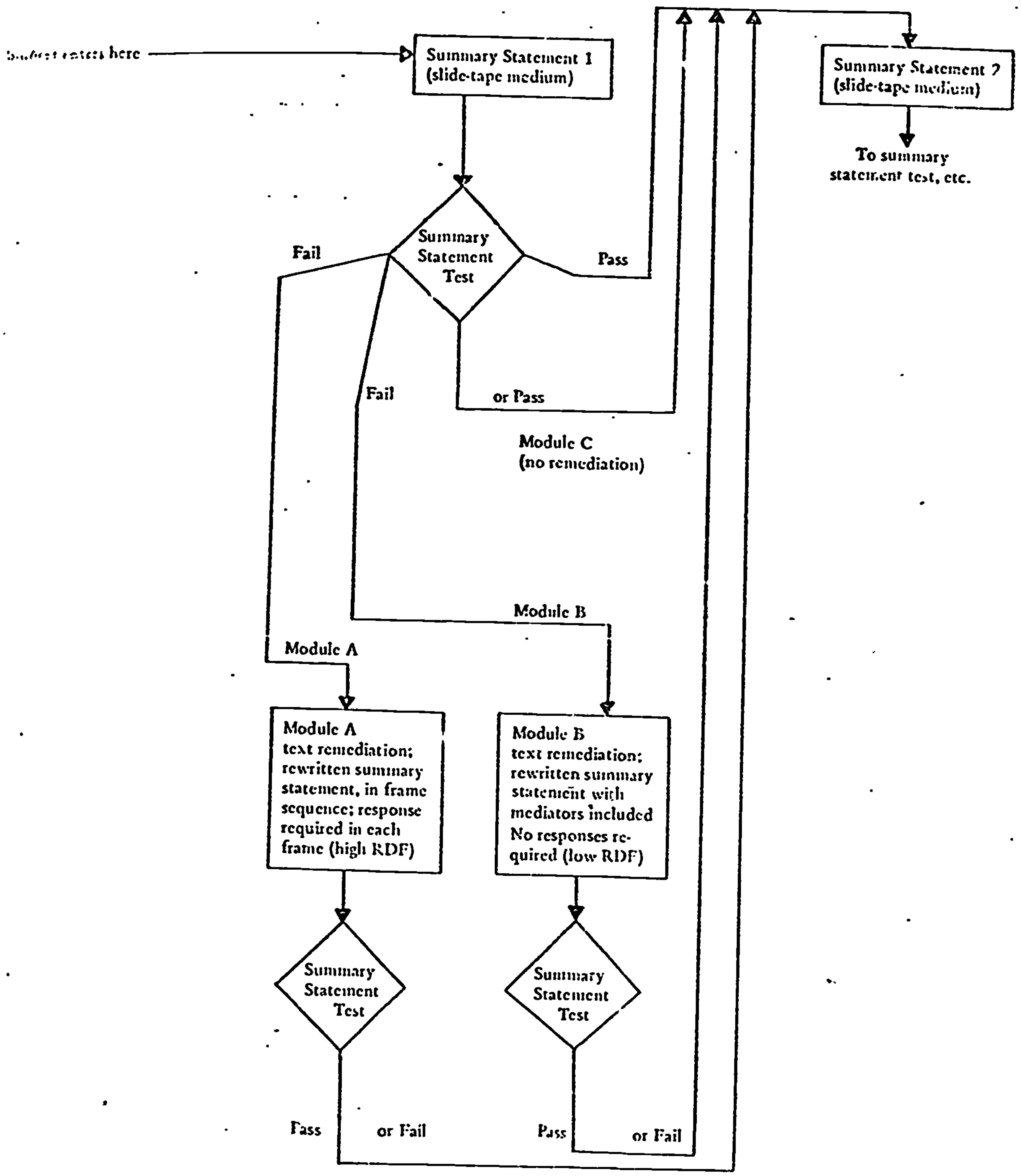

Figure 16. Illustration of Experimental Procedure for Design of Hypothesis IV. Example Flowchart of Instruction Sequence of CPT Unit 2.

An important feature of this design is that within each CPT Unit, there are three experimental treatments and six groups (rows) of students. Therefore, two groups of students combine to take each treatment. The purpose of this arrangement is to achieve complete counterbalancing of treatment sequences (in two orthogonal squares), so that (1) no single student will suffer inordinately from one of the experimental conditions, and (2) so that the evaluation of both components of the units X treatments interaction between students is permitted.

4.2.4.3 Statistical Procedures. The analysis of variance source table for the statistical evaluation of Hypothesis IV is shown in Figure 17 below.

Source	Formula	df	Error Term
Squares (Q)	t-2	1	
Row (R/Q "UT")	$(t-1)^2$	4	S/RQ
Students (S/RQ)	$(t-1)^2$ (n-1)	42	
Units (U)	t-1	2	S X (UT)/R
Treatments (T)	t-1	2	S X (UT)/R
Residual "UT"	$(t-1)^2$	4	S X (UT)/R
U X Q	t-1	2	S X (UT)/R
T X Q	t-1	2	
S X (UT)/R	t $(t-1)^2$ (n-1)	84	
Total	(t-2) (t-1) (n)-1	143	

Figure 17. Analysis of Variance Source Table (simplified) for the Statistical Evaluation of Hypothesis IV.

There is one additional note related to the design and statistical procedure: The T X Q and U X Q interactions permit evaluation of carry-over effects. This is particularly important in the present hypothesis because of predicted differences between remediation procedures.

4.2.5 Hypothesis V

> Using "peer-monitor" procedures will significantly improve student performance in an instructional sequence.

4.2.5.1 Introduction. In each of the four previous hypotheses, designed to assess the effects of presentation variables, the direction and control of the instruction is largely external. Although the students proceed through the course at their own pace, the management decisions in particular modules are predetermined and embedded within the presentation. Based on his performance, a student is instructed to read this or that remediation material, go to one or another enrichment sequence, progress to the next segment or repeat the first one, etc. Essentially, student progress in all of these activities is monitored by the instructional system.

A number of investigators, however, have suggested that instructional procedures may be significantly improved if the student is allowed to participate in his own management. As Tosti (1968) has written:

> Another and often overlooked approach is to get the student to do more of his own management. This has sometimes been characterized as making the student an "independent learner"

(although there may be other behaviors implied by this phrase). Even though he is still dependent on the system, he could be allowed to make more decisions on his own, which may be of more value than just solving the logistic problems. In a recent experiment with junior college students, in an individualized program, self-management was demonstrated to be the most significant factor in student success (Csanyi and Beck, 1968).

Another monitoring form, once common in the one-room schoolhouse, but now mostly discontinued, is peer management. Students may be divided into pairs of teams in which the students manage each other's instructional presentation.

Keller (1966) and others have recently extended the concept of peer-management into an elaborate system of individualized instruction in which the student manages a great deal of his behavior in the course.

Keller (1966) has described his basic procedures, applied in the teaching of an introductory psychology class:

> The unique features of our method can be presented best, perhaps, by quoting from a mimeographed hand-out which is given to each student on his first day of class-attendance:
>
> This is a course that permits you to move, from start to finish, at your own pace. You will not be held back by other students or forced to go ahead until you are ready. At best, you may meet all the requirements in less than one semester; at worst, you may not complete the job within that time. How fast you go is up to you.
>
> The work of this course will be divided into certain units of content, which correspond roughly to a series of homework assignments and laboratory exercises. These units will come in a definite numerical order, and you must show your complete mastery of each unit (by passing a readiness test) before moving on to the next.
>
> A good share of your reading may be done in the classroom, at those times when no lecture, demonstrations, or other activities are taking place. Your classroom, that is, will sometimes be a study-hall.

The lectures and demonstrations in this course will have a different relation to the rest of your work than is usually the rule. They will be provided only when you have demonstrated your readiness to appreciate them; no examination will be based upon them; and you need not attend them if you do not wish. When a certain percentage of the class has reached a certain point in the course, a lecture or a demonstration will be available at a stated time, but your presence there will not be compulsory....

You will have an opportunity to ask questions and to discuss the course with your classmates, your proctor, your course assistant, or your instructor. Group discussion will depend not only upon your desire, but also upon your preparation --- your readiress --- to engage in such activity. Questions may be asked in writing at any time, and will be individually considered by your instructor.

The teaching staff of your course will include, as already suggested, a proctor, a laboratory assistant, a classroom assistant, and your instructor. The proctor has been chosen for his mastery of the course material and orientation, for his maturity of judgment, for his understanding of the special problems that confront you as a beginner, and for his willingness to assist. He will provide you with all your study materials except your textbooks. He will ... pass upon your readiness tests as satisfactory or unsatisfactory. His judgment will ordinarily be law, but if he is ever in serious doubt he can appeal to the classroom assistant for a ruling, or even to the instructor. Failure to pass a test on the first try, the second, the third, or even later, will not be held against you; better too much testing than not enough, if your final success is to be assured....

Your work in the laboratory will be carried out under the direction of a graduate laboratory assistant, whose detailed duties need not be listed here. In addition, there will be a graduate classroom assistant, upon whom your proctor will depend for various course materials (assignments, study questions, special readings, and so on) and who will collect and keep up to date all progress records for all course members. The classroom assistant will confer with the instructor daily, aid the proctors on occasion, and act in a variety of ways to further the smooth operation of the course machinery.

The instructor will have as his principal responsibilities: (a) the selection of all study materials used in this course; (b) the organization and the mode of presentation of these materials; (c) the construction of tests; and (d) the final evaluation of each student's progress...It will be his duty also to provide lectures, demonstrations, and discussion opportunities for all students who have earned the privilege; to act as a clearing-house for requests and complaints; and to arbitrate in any case of disagreement or misunderstanding between students and proctors or assistants.

All students in the course are expected to take a final examination, in which the entire term's work will be represented. With certain.... exceptions to be mentioned later, this examination will come at the same time for all students, at the end of the term. [The exceptions were those students who finished early and chose to be examined immediately, and those students who took a grade of Incomplete.] The examination will consist of questions which, in large part, you have already answered in your readiness tests. Twenty-five percent of your course grade will be based upon this examination; the remaining seventy-five percent of your grade will be based upon the number of units of reading and laboratory work that you have completed successfully during the term."

Along with this description and a few instructions, the student was given his first assignment and his course began. The main classroom was used as a study hall; a smaller classroom was used for test-taking; and a third room, containing seven proctors' cubicles, was set aside for grading and discussion of tests. When a student in the study hall felt ready for examination, he went to the test room, where the assistant gave him a test form and his blue book. When the questions were answered, the student returned the test form and took his blue book to his proctor. After the test was evaluated, it was retained by the proctor, and the student either received a new assignment or was sent back to study for another test on the same material, usually after considerable discussion of the errors made. The graded test was passed on to the classroom assistant or the instructor at the end of the hour, to be checked through and recorded.

It should be noted that the first proctors were "volunteers from the top 20 of a conventional lecture course having the same general content as (the experimental) class." Later on, "non-majors, sophomores, and, in one instance, a second term freshman" served as proctors. "A well motivated A-student," Keller suggested, "without obvious personality defects, and under adequate supervision, is likely to met the requirements of the job, without serious trouble. Moreover, Keller continued,

> ...I am at this time convinced that, with proper supervision and moderately careful screening, for the task at hand, these proctors constitute one of the best features of our course --- if not the best. They are the ones who provide, for each individual student, that personal contact which is commonly denied to all but a select few in most of our classes, and which is especially meaningful for the beginner in any field. For this beginner, the instructor, and even the assistant, may be a far-off figure in another world, but the proctor is a person of status, who has the answers, understands the problems, and is willing to help. The relation between the student and the proctor is established more firmly and workably with every test that is graded --- whether it be a pass or fail. And every step forward made by the student is almost as satisfying to the proctor as it is to him. A better system of mutual reinforcement is hard to picture.

Since it appears that the proctor-student interaction may be the most important characteristic in this system, it is only a procedural alteration to use as proctors those students in the course who have completed a particular segment of instruction. The proctors would then be the peers of the students who are monitored; hence, the term "peer-monitor system."

As an example of how the system would operate, consider two students progressing at approximately the same pace through the.

Leadership Management Course. In a particular segment, both would be given material to study, covering the content of that segment, and would be requested to report to the instructor when they felt they were ready for the progress check. The first student to report would be given an alternate form of the progress check. He would be asked to take the test, correct any errors, and be prepared to quiz the second student on the progress check when that student was ready. The first student would thus be a peer-monitor for the second. After the interaction between the students, both would take the final progress check. Students in a control condition would be given the alternate progress check and asked to correct any errors themselves (without consulting another student) and to return when ready for the final progress check. It is predicted that the students involved in the peer-monitor interaction will perform significantly better on the progress checks and CPT Tests than those students not participating in the interaction.

4.2.5.2 Design and Procedure. Because the test of Hypothesis V is planned for the second half of the Leadership Management Course, formal specification of the design and statistical procedures must await an analysis of the content and logistics of those CPT Units. It is expected, however, that the treatment groups will be compared in a within-subjects design. Each student will experience each of the three following conditions:

Module A -- student acts as peer-monitor

Module B -- student is monitored

Module C (control) -- student does not participate in the peer-monitor system

Different groups of students will experience the experimental treatments in a completely counterbalanced order, as in Hypothesis III. Since three CPT Units will not be available, the progress check will be a primary measure of performance. However, one CPT Unit may be designed so that some groups will be given Modules A and B, and the other students given Module C. In this design, the CPT Test scores may be used to assess the differences within the unit between students who participate in the peer-monitor procedures and those who do not.

5. CONCLUDING REMARKS

Perhaps the most appropriate way of concluding this description of the research and evaluation plans is to reproduce, with a few slight modifications, the final paragraphs of Tosti and Ball's (1968) paper on the media selection model, which discusses the larger overall considerations and goals of continuing media research:

> It would be comforting to think that the engineering model proposed in this report would appear immediately in operational educational systems and revive the ed' bizz'.
>
> However, the model proposed here is neither sacred nor sufficient.
>
> An initial look at the dimensions of media produced a chart containing over 100 dimensional candidates organizable into small numbers of generic dimensions only with great latitude in the meaning assigned to generic healings. Many of the dimensions were concerned with general learning theory and threatened to broaden the media selection model beyond hope of a practical solution.
>
> The present.....dimensions of presentation design are a good first attempt to structure presentation in terms meaningful to the behavioral engineer. Of course, the underlying philosophy in these dimensions and the whole model is that behavior engineering is a practical solution to the needs of education. Overt considerations of behavior engineering are foreign to present educational practice, but Homme and others feel that an acknowledgement of the efficacy of behavioral engineering and its dedicated practice in the schools will solve most of the educators' problems.
>
> Presentation engineering was discussed in terms of.....empirical dimensions. A continuation into procedures for selecting presentation form must be made. Any real

situation will also be finally judged in the marketplace where cost, equipment, transition, and completeness will be prime variables. Much work is needed to devise and measure alternate systems which will be responsive to the demands of the market. The creation of demand for a fixed product is another strategy which the diversity of educators' interests, goals and experience and their autonomy seems to preclude.

Media limitations is another research area which will ultimately produce media selection criteria together with behavioral change expectancy figures.

If media preferences really exist and if learning rates of individuals are sensitive to media choice, then some measurable student attributes must be defined which will provide media selection guides. It would be optimal to be able to derive curves for expected behavioral change versus media selection on a real-time basis for each student, and so provide material to achieve highest expected gain.

A more practical solution may be to research thoroughly the techniques of behavioral engineering of the student so that a reasonable (but possibly not optimal) gain is achieved with a practical media multiplicity. Not enough is yet known about media cues which promote media preference ' engineer the presentation over several media devices. The domain of presentational and media cues must be explored in depth.

Finally, the complexities of creating and maintaining an operational system must be sorted out. The job of educating will be only half done when a system is on the store shelf. Success will be largely a function of administering a smooth operating system that works in the field, cranking out the required student goals as was predicted on paper.

REFERENCES

Allen, H. Research in instructional media and art education. Final Report of the Uses of Newer Media in Art Education Project. National Art Education Association, Washington, D.C., August 1966.

Alter, M., and Silverman, R. The response in programmed instruction. Journal of Programmed Instruction. Journal of Programmed Instruction. 1962, 1 (1), 55-68.

Ash, P., and Jaspen, N. The effects and interactions of rate of development, repetition, participation, and room illumination on learning from a rear projected film. Technical Report SDC 269-7-39. Instructional Film Research Reports. Port Washington, Long Island, N.Y.: U.S. Naval Special Devices Center, 1953.

Becker, S.L., et. al. Communication skills. An experiment in instructional methods. Iowa City: The State University of Iowa, August 1958.

Berelson, B. and Steiner, G.A. Human Behavior: An Inventory of Scientific Findings. New York: Harcourt, Brace and World, Inc., 1964.

Berger, R. J. A test of the Premack hypothesis with responses to audio and visual stimuli. Unpublished doctoral dissertation, Arizona State University, 1965.

Bloom, B.S. (Ed.) Taxonomy of Educational Objectives. New York: Longmans, Green, 1954.

Briggs, L.J. Self-pacing versus automatic pacing in paired-associate learning. In Lumsdaine, A.A. (Ed.) Student response in programmed instruction. A symposium on experimental studies of cue and response factors in group and individual learning from instructional media. Sponsored by Headquarters, Air Research and Development Command, U.S. Air Force. Washington, D.C.: National Academy of Sciences, National Research Council, 1961.

Brown, J.I. The importance of structured outside reading assignments. College and Adult Reading, II, Second Annual Yearbook, North Central Reading Association, 1963. (Abstracted in part in NAEB Journal 21, March 1962.

Buros, O.K. The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1959.

Buros, O.K. The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.

C'de Baca, J., and Chadwick, C.B. PROMOD: A new approach to multi-level programming. NSPI Journal, 1968, 7 (2), 8-10.

Carpenter, C.R., and Greenhill, C.P. Comparative Research on Methods and Media for Presenting Programmed Courses in Mathematics and English. University Park: Pennsylvania State University, March 1963.

Csanyi, A.P. An investigation of visual versus auditory programming in teaching foreign language pronunciation. In Investigations of the Characteristics of Programmed Learning Sequences. Pittsburgh: University of Pittsburgh, Programmed Learning Laboratory, 1961. 135-150.

Csanyi, A.P. Analytical techniques for programed instruction in foreign languages. Paper presented at the Annual Convention of the National Society for Programmed Instruction. May, 1965.

Csanyi, A.P., and Beck, M.H. Final Report on Phase I: Junior College Demonstration Project. Albuquerque: Behavior Systems Division, Westinghouse Learning Corporation, July 31, 1968.

Cronbach, L. J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.

Crowder, N.A. Automatic tutoring by means of intrinsic programming. In Lumsdaine, A.A., and Glaser, R. (Eds.). Teaching Machines and Programmed Learning. Washington: National Education Association, 1960, 286-298.

Educational Testing Service. Designing Validity Studies and Collecting Data. College Entrance Examination Board, 1967.

Edwards, A.L. Experimental Design in Psychological Research. New York: Holt, Rinehart and Winston, 1950.

Eigen, L.D. High school student reaction to programmed instruction. Phi Delta Kappan, 1963, 44, 282-285.

Ellis, H.C. Research design in programmed learning. National Society for Programmed Instruction. August 1962, 7-8.

Evans, J. Behavioral objectives are no damn good. Technology and Innovations in Education. New York: Frederick A. Praeger, 1968.

Evans, R.I. An examination of students' attitudes toward television as a medium of instruction in a psychology course. Journal of Applied Psychology, 1956, 40, 32-34.

Feldhusen, J.F., and Birt, A. A study of nine methods of presentation of programmed learning material. Journal of Educational Research, 1962, 55, 461-466.

Feldhusen, J.F., and Eigen, L.D. Interrelationships among attitude, achievement, reading, intelligence, and transfer variables in programed instruction. Paper presented to the Midwestern Psychological Association, May, 1963.

Ferster, C.B., and Sapon, S.M. An application of recent developments in psychology to the teaching of German. Harvard Educational Review, 1958, 28, 58-69.

Follettie, J.F. Effects of training response mode, test form, and measure on acquisition of semi-ordered factual materials. Research Memorandum 24. Fort Benning, Georgia: U.S. Army Infantry Human Research Unit, April 1961.

Frank, J. H. An evaluation of closed circuit television for interceptor pilot training. Dissertation Abstracts, 1955, 15, 2060-2061.

Frye, C.H. Group vs. individual pacing in programmed instruction. In Schramm, W. The Research on Programmed Instruction. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1964. (OE-34034) (Abstract).

Fulton, W.R., and Rupiper, O.J. Selected vicarious experiences versus direct observational experiences of pre-service teachers in the foundation areas of professional preparation at the University of Oklahoma. USOE Project No. 192. Norman: College of Education, University of Oklahoma, 1961.

Gagne, R.M. The Conditions of Learning. New York: Holt, Rinehart and Winston, 1965.

Gagne, R.M., and Brown, L.T. Some factors in the programming of conceptual learning. Journal of Experimental Pyschology, 1961, 62, 313-321.

Gilpin, J. Design and evaluation of instructional systems. Paper presented in a symposium at the 1961 meetings of the American Psychological Association. Richmond, Indiana: Earlham College, 1961.

Glaser, R. and Schaefer, H. Principles of Programming Printed Materials. Pittsburgh: University of Pittsburgh, 1961. Final report on Contract AF 33(616)-7175.

Goldman, L. Using Tests in Counseling. New York: Appleton-Century-Crofts, Inc. 1961.

Goldstein, L.S., and Gotkin, L.G. A review of research: teaching machines versus programmed textbooks as presentation modes. Journal of Programmed Instruction, 1962, 1 (1), 29-36.

Greenhill, L. P. Review of trends in research on instructional television and film. In: Reid, J.C., and MacLennan, P.W. Research in Instructional Television and Film. U.S. Department of Health, Education, and Welfare, Bureau of Research. Catalog #FS-5.231:34041, 1967.

Gropper, G. L. Learning from Visuals: Some Behavioral Considerations. AV Communication Review, Spring, 1966.

Gropper, G. L. Why is a picture worth a thousand words? AV Communication Review, 1963, 11, 75-95.

Gropper, G. L. and Lumsdaine, A.A. An experimental comparison of a conventional TV lesson with a programmed TV lesson requiring active student response. Studies in Televised Instruction, Report No. 2. Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX, and American Institute for Research, 1961a.

Gryde, S.K. The feasibility of programmed television instruction. AV Communication Review, 1966, 14 (2), 71-89.

Gulliksen, H. Theory of Mental Tests. New York: John Wiley & Sons, Inc., 1950.

Hatch, R.S. and Flint, L.L. Programed Learning: A Comparative Evaluation of Student Performance Variables Under Combinations of Conventional and Automated Instruction. New York: U.S. Industries, Educational Sciences Division, 1962.

Hickey, A.E., Autor, S.M., and Robinson, E.J. Programmed Instruction Integrated with Broadcast ETV. Boston: Northeastern University, 1962.

Holland, J.G. Evaluating teaching machines and programs. Teachers' College Record, October 1961, 56-65.

Holland, J.G. Teaching psychology by a teaching machine program. In Berelson, B. and Steiner, G.A. Human Behavior: An Inventory of Scientific Findings. New York: Harcourt, Brace and World, Inc., 1964.

Homme, L.E. (principal investigator). A demonstration of the use of self-instruction and other teaching techniques for remedial instruction of low-achieving adolescents in reading and mathematics. Submitted to U.S. Office of Education, August 1965.

Homme, L.E. Contiquity theory and contingency management. Psychological Record, 1966, 16 (3), 233-241.

Homme, L.E., and Tosti, D.T. Contingency management and motivation. NSPI Journal, 1965, 4 (7), 14-16.

Kasten, D.F., and Seibert, W.F. A study of televised military science instruction. TVPR Report No. 9. Lafayette, Indiana: Purdue University, July 1959.

Keller, F.S. Engineering individualized instruction in classroom teaching. Paper presented at the Rocky Mountain Psychological Association in Albuquerque, New Mexico, May 13, 1966.

Kimble, G.A., and Wulff, J.J. The effects of 'response guidance' on the value of audience participation in training film instruction. Human Factors Operations Research Laboratories, USAF, Report No. 35. Audio-Visual Communication Review, 1953, 1, 292-293. (Abstract).

Kring, W.K., and Stolurow, L.M. Predictive Variables. Technical Recommendations No. 5.2 USNA Contract No. N00161-7339-4781, July 25, 1968.

Levitt, E. Clinical Research Design and Analysis in the Behavioral Sciences. Springfield, Illinois: Charles C. Thomas (publisher), 1961.

Lindquist, E.F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Company, 1956.

Logan, F.A. Micromolar behavior theory and performance speed in education. Harvard Educational Review, 1963, 33(2), 178-185.

Loree, M. R. Psychology of Education. New York: Ronald Press, 1965.

Lyman, H.B. Test Scores and What They Mean. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Maccoby, N. and Sheffield, F.D. Theory and experimental research on the teaching of complex sequential procedures by alternate demonstration and practice. In symposium on Air Force human engineering, personnel, and training research. Washington, D.C.: National Academy of Sciences National Research Council, 1958.

Macomber, F., et. al. Experimental Study in Instructional Procedures. Oxford: Miami University, October 1, 1957.

Mainland, D., and Murray, I.M. Tables for use in fourfold contingency tests. Science, 1952, 116, 591-594.

Margolius, G.J. and Sheffield, F.D. Optimum methods of combining practice with filmed demonstration in teaching complex response sequences: Serial learning of a mechanical assembly task. In: Lumsdaine, A.A., (ed.), Student Response in Programed Instruction: A Symposium on Experimental Studies of Cue and Response Factors in Group and Individual Learning from Instructional Media. Washington, D.C.: National Academy of Sciences, National Research Council, 1961.

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

Moore, J.W., and Smith, W.I. Knowledge of results in self-teaching spelling. Psychological Reports, 1961, 9, 717-726.

Parry, S.B. To err is human ... and sometimes desirable. Programed Instruction Newsletter, 1963, 2 (4), 4-5.

Porter, D. Some effects of year long teaching machines instruction. In E. Galanter (ed.), Automatic Teaching. New York: John Wiley & Sons Inc. 85-102.

Reid, J.C. and MacLennan, P.W. Research in Instructional Television and Film. U.S. Department of Health, Education, and Welfare, Bureau of Research, Catalog #FS-5.234:34041, 1967.

Rosenthal, R. Experimenter Effects in Behavioral Research. New York: Appleton-Century-Crofts, 1966.

Schramm, W. The Research on Programed Instruction. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1964. (OE-34034)

Shay, C.B. Relationship of intelligence to step size on a teaching machine program. Journal of Educational Psychology, 1961, 52, 98-103.

Silberman, H.F. Self-teaching devices and programed materials. Review of Educational Research, 32, 179-193.

Silverman, R., and Alter, M. Response mode, pacing, and motivational effects in teaching machines. Port Washington, N.Y.: U.S. Naval Training Devices Center. Technical Report 507-03. 1961.

Smith, R.G. Jr. The Design of Instructional Systems. The George Washington University Human Resources Research Office. Technical Report 66-18. November 1966.

Smith, W.I. and Moore, J.W. Programed Materials in Mathematics for Superior Students in Rural Schools. Lewisburg, Pa.: Bucknell University, 1962.

Stolurow, L.M. Instructional systems validation planning. Technical recommendation No. 2. USNA Contract No. N00161-7339-4781, March 29, 1968.

Stolurow, L.M. Implications of current research and future trends. Journal of Educational Research, 1962, 55, 518-27.

Taber, I.I., Glaser, R., and Schaefer, H.H. Learning and Programmed Instruction. Massachusetts: Addison Wesley, 1965.

Tosti, D. T., and Ball, J.R. A Behavioral Approach to Instructional Design and Media Selection. Albuquerque, 1968. Behavior Systems Division, Westinghouse Learning Corporation.

Tosti, D.T. Muddled media. Paper presented at the American Association for the Advancement of Science, May 27, 1968.

Twyford, L. Film profiles. Instructional Film Research Reports. Technical Report SDC 269-7-23. Port Washington, Long Island, N.Y.: U.S. Special Devices Center, 1951.

Van Atta, L. Behavior in Small Steps. Contemporary Psychology, 1961, 10, 378-81.

Weiss, W., Maccoby, N., and Sheffield, F.D. Combining practice with demonstration in teaching complex sequences: Serial learning of a geometric-construction task. In Lumsdaine, A.A., (ed.), Student Response in Programed Instruction: A Symposium on Experimental Studies of Cue and Response Factors in Group and Individual Learning from Instructional Media. Washington, D.C.: National Academy of Sciences, National Research Council, 1961.

Westover, F.L. A comparison of reading and listening as a means of testing. Journal of Educational Research., 1958, 52, 23-26.

Wood, D. A. Test Construction. Columbus, Ohio: Charles E. Merrill Books, Inc., 1961.

APPENDIX

NOTES ON THE MULTIPLE-REGRESSION STATISTICS TO BE USED IN EVALUATING THE RELATIONSHIPS BETWEEN STUDENT CHARACTERISTICS AND PERFORMANCE

The sources of variance used as error terms in the analysis of variance represent unexplained individual differences which will be subjected to additional analysis to identify subject variables related to learning achievement. In the analysis of Figure 7, the variance among subjects is computed from the deviations of each student's mean score from the mean of his own group. Regression of these deviation scores on selected subject variables will identify those variables associated with individual differences in overall performance during the experiment. Variation in the regression for subjects under different row treatments will identify variables associated with individual differences in response to treatments, i.e., subject X row-treatment interactions. The subject X unit variance comes from deviations of each student's scores from the means of the Latin Square cells and from the student's overall mean, and may be similarly analyzed to identify variables associated with individual differences in response to counterbalanced treatments.

In addition to studying single regressions for subject variables where specific a priori hypotheses may be set up, it is desirable for interpretive purposes to obtain multiple regressions predicting as much of the error variances as possible. As is well known, single regressions may be misleading indications of the "true" importance of a variable in prediction, since the contribution of other correlated predictor variables is not partialed.

out (Darlington, 1968). However, considering the small number of degrees of freedom available in the error sources, the number of variables which can be simultaneously entered into regression is severely limited.

Therefore, the regression analysis will be conducted in two stages. In the first stage, multiple regressions will be developed separately for data under each treatment condition, using the stepwise method for selecting predictor variables from the entire set available, as described by Elfroymson (1962). On the basis of study of the results of the first stage, a much smaller number of important predictors will be retained in a second multiple regression analysis performed on the combined data.

In addition to first-order linear terms, second order quadratic terms will be included in the regression model as a check on nonlinearity of regression, as well as product terms to check possible interactions among the subject variables. Assuming the selection of K = 3 variables in the first stage, the source table for the complete second stage analysis is shown in Figure 18. The contribution of each of the K variables to the sources shown in Figure 18 would, of course, be examined separately in the actual analysis, but are combined here for brevity. The final form of the analysis may also be considerably reduced by pooling with the deviation sources those components of variance which show little or no contribution to prediction. Quadratic or product terms may also be eliminated on the basis of the results of the analysis in the first stage. Final tests of significance would be F-tests, using the pooled deviation sources as estimates of pure error.

Source	df Formula	df
Subjects S/RM	mr (n - 1)	44
Regression	$k^2 + 3k$	18
Linear (L)	k	3
Quadratic (Q)	k	3
Product (P)	$\frac{k(k-1)}{2}$	3
L X M	k	3
Q X M	k	3
P X M	$\frac{k(k-1)}{2}$	3
Deviation from	mr (n - 1) — k (k - 3)	26
S X U/RM	mr (n - 1) (t - 1)	132
Regression		108
L X Units	3k	9
L X Treatments	5k	15
L X Residual	4k	12
Q X Units	3k	9
Q X Treatments	5k	15
Q X Residual	4k	12
P X Units	$3\frac{k(k-1)}{2}$	9
P X Treatments	$5\frac{k(k-1)}{2}$	15
P X Residual	$4\frac{k(k-1)}{2}$	12
Deviation from Reg.	mr (n - 1) (t - 1) — 6k (k + 3)	24

Figure 18 -- Source Table for Second Stage of the Multiple Regression Analysis

APPENDIX REFERENCES

Darlington, R.B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

Elfroymson, R.A. Multiple regression analysis. In N.R. Draper and H. Smith (eds.), Applied Regression Analysis. New York: Wiley, 1966.