DOCUMENT RESUME

ED 071 345 EM 010 490

AUTHOR	Bessemer, David W.; Rivers, LeRoy C.
TITLE	Introduction to Psychology and Leadership. Report of Phase II Research Findings. The Design and Methodology for Research on the Interaction of Media, Conditions of Instruction, and Student Characteristics for a Multimedia Course in Leadership, Psychology and Management. Part I: Conditions of Instruction.
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-6-12a
BUREAU NO	BR-8-0448
PUB DATE	15 Oct 70
CONTRACT	N00600-68-C-1525
NOTE	209p.; See also EM 010 418 and EM 010 419
EDRS PRICE	MF-$0.65 HC-$9.87
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

This report is the first of a two-part document which presents the results of the research conducted during the first implementation of the leadership, psychology and management course at the United States Naval Academy. A detailed discussion of the background of the research, the research plan, and its implementation is provided. The second part of this document is EM 010 491, the final report is EM 010 418, EM 010 419, and EM 010 484, and EM 010 418 through EM 010 447 and EM 010 451 through EM 010 512 are related documents. (Author/RH)

Set #3

Westinghouse Learning Corporation

Contract No. N00600-68-C-1525

REPORT OF PHASE II RESEARCH FINDINGS: THE DESIGN AND METHODOLOGY FOR RESEARCH ON THE INTERACTION OF MEDIA, CONDITIONS OF INSTRUCTION, AND STUDENT CHARACTERISTICS FOR A MULTIMEDIA COURSE IN LEADERSHIP, PSYCHOLOGY AND MANAGEMENT

PART I: CONDITIONS OF INSTRUCTION

TR-6.12a October 15, 1970

TR-6.12a
Oct., 1970

REPORT OF PHASE II RESEARCH FINDINGS: THE DESIGN AND METHODOLOGY FOR RESEARCH ON THE INTERACTION OF MEDIA, CONDITIONS OF INSTRUCTION, AND STUDENT CHARACTERISTICS FOR A MULTI-MEDIA COURSE IN LEADERSHIP, PSYCHOLOGY AND MANAGEMENT

PART I: CONDITIONS OF INSTRUCTION

Contract No. N00600-68-C-1525

October, 1970

ABSTRACT

This report is the first of a two-part document which presents the results of the research conducted during the first implementation of the Leadership, Psychology and Management course at the United States Naval Academy. A detailed discussion of the background of the research, the research plan, and its implementation is provided.

Prepared by

David W. Bessemer
LeRoy C. Rivers

Approved by

Project Manager
Leadership Management Course

WESTINGHOUSE LEARNING CORPORATION
2083 WEST STREET
ANNAPOLIS, MARYLAND 21401

TABLE OF CONTENTS

APPENDICES

TABLES

I. INTRODUCTION

This report describes the results of several studies conducted in the experimental multimedia Leadership Management course developed by Westinghouse Learning Corporation (WLC) for the United States Naval Academy (USNA). These studies form part of a comprehensive investigation of factors influencing student achievement intended to guide continuing improvement of the Leadership Management course, and to have wide-ranging application in the field of educational technology.

The research presented in Part One of the report was designed to evaluate the effects of major variations in conditions of instruction involving media and presentation forms. Tests of five specific hypotheses were conducted with effects of experimental manipulations measured by three types of tests reflecting accomplishment of three broadly different kinds of learning tasks. The relationship of student learning in specific conditions of instruction to individual characteristics of the student is the subject of Part Two of this report. Later reports will deal with additional research on major operational organization variables relevant to the final design and implementation of the total instructional process in the Leadership Management course.

The WLC research plan has the distinction of being one of the first to provide a joint examination of factors in all major categories relevant to the design of an instructional system, including media, presentation forms, task requirements, student characteristics, and operational organization. The WLC plan is unique both in the number of factors investigated and in the use of an entire ongoing course system as an experimental vehicle permitting empirical findings to be extracted relevant to the influence of each factor singly and in combination. It is expected that experiments of this type, as part of a concentrated effort in educational research, may 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.

II. BACKGROUND OF THE RESEARCH

In this section an informal analysis of educational systems is presented leading to the identification of major classes of variables important in the development of a comprehensive educational technology. An important distinction between presentation and media variables is then described in relation to conceptual organization of these variable classes based on Tosti and Ball's (1969) instructional design and media selection model. Implications of some fundamental methodological considerations in the development of presentation and media research are discussed. The section concludes with a brief review of task, student, and operational variables in terms of their relevance to the WLC research plan.

Variables in Instructional Systems

Instructional systems may be analyzed as an interactive process among components of (1) designed behavioral objectives, (2) student, (3) materials designed to change student behavior toward the designed objectives, (4) media presenting materials to the student, and (5) operational organization bringing these components into articulation. The nature of such a system is illustrated in Figure 1.

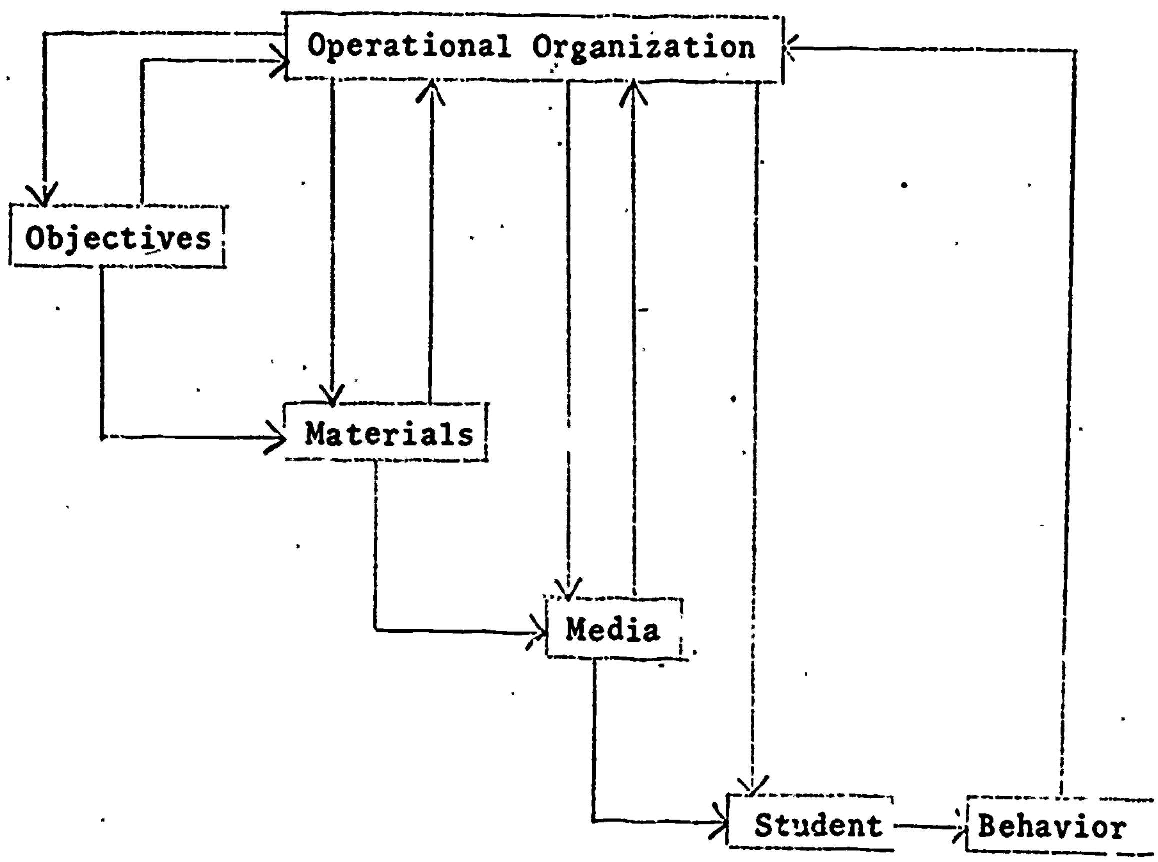


Figure 1. Components of educational system.

An analysis of this kind clarifies several points in relation to instructional systems. First, evaluation of a system rests primarily on examination of the correspondence between objectives desired and behavioral changes achieved. Other criteria, such as the desirability of objectives and cost-effectiveness are external to the system and concern the utility of the system as a component in larger social systems. Second, research on instructional systems requires the manipulation or measurement of characteristics (variables) differing among individual examples of the same component, and study of the associated modifications in the behavioral output. Clearly, evaluation of a standing instructional system may proceed, given information on objectives and behavior change, but when a discrepancy is discovered between these, efforts must return to the research domain to discover how to modify the system to eliminate the discrepancy.

Corresponding to each component of the instructional system are a large number of variables potentially important in determining the final behavior change effected. The instructional system designer is, however, not equally free to manipulate or select settings for the variables in all categories. Once objectives for a given target population of students are established for the system, the range of variation in task variables is fixed by the objectives, and student variables are fixed by the definition of the target population. The system designer must then work with the presentation variables of the materials, media variables, and operational variables to arrange an optimally effective system for the particular types of tasks and students involved. From this point of view, presentation, media, and operational variables are of primary research interest, while task and student variables are of interest mainly in relationship to the other catgories of variables.

Informal analyses of this type have provided heuristic guidelines in the development of WLC's research plan, and should prove useful in similar efforts in the future.

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, film, cartridge tape, and computerized instruction) there is a steadily increasing variety of devices available for use in any instructional program. The problem is to decide what medium is best for a particular purpose, and how to efficiently utilize its capabilities.

In addressing this problem, Tosti and Ball (1969) have developed a model for instructional 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) necesary 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 instructional design model [illegible] 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 sequence by identifying its characteristics along basic dimensions which are common to all instructional presentations. Since the specification of presentational variables is a critical consideration in educational research, the application of Tosti and Ball's model may result in a significant improvement in the quality of studies in educational technology 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 application of the Tosti and Ball model involves the determination of a precise presentation design for each instructional objective. Media are then selected on the basis of their limitations in presenting the presentation design intact. The primary question raised by the Tosti and Ball model is whether variations of conditions of instruction in the presentation design domain are of greater or lesser importance than variation in the media domain with respect to student achievement. The implication is that if the presentation design is held constant over a unit of instruction the use of different media should not result in differential levels of achievement. Conversely variations in the presentation design with the medium constant over a unit of instruction should result in differential levels of achievement.

It would appear that a detailed behavioral analysis of individual learning events would require the specification of a host of subsidiary variables associated with each of the Tosti and Ball presentation variables. The general approach used here was to investigate the presentation design--media controversy over larger units of instruction involving similar types of instructional objectives. The research planned for the Leadership Management course may be considered a test of the utility of the Tosti and Ball instructional design model as it pertains to their presentation design variables. In addition, the research may indicate whether or not experimental analysis at the macro-level (over large units of instruction) can identify major pervasive influences on instructional effectiveness. If this type is the case, it may be possible to sidestep the much more detailed and expensive analysis required at the micor-level of instruction. This would allow for the development of a more simplified and economic form of instruction design than would otherwise be possible. A detailed description of the research plan is given in Chapter III.

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 of Tosti and Ball's model. The dimensions of presentation are discussed in detail in the following section.

<u>Dimensions</u> <u>of</u> <u>presentation</u>.

presentation have been derived by a logical analysis of instructional systems (Tosti and Ball, 1969). 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)

Each capability may be further differentiated in terms of two attributes: form and frequency. The result of this analysis is a 3 x 2 matrix, represented in Table 1, in which six dimensions of presentation are generated. Further study of common or possible instructional procedures reveals a number of levels or categories associated with each dimension, also listed in Table 1.

The following subsections give a detailed description of each dimension and its levels or categories. Discussions of research findings relating to the dimensions will be presented in connection with the experiments involving those dimensions.

Stimulus form (representation). This dimension is related most directly to media. It characterizes the dominant mode of sensory reception (by the student) of the instructional material, inherent in the means of representation of stimuli. 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

TABLE I

SYSTEM CAPABILITY MATRIX AND

DIMENSIONS OF PRESENTATION[a]

System Capability	Attributes of System Capability	
	Form	Frequency
Stimulus	<u>Stimulus Representation</u> Verbal-written Verbal-spoken Pictorial	<u>Duration</u> Transient-Persistant Length of time the presentation remains intact a. low b. intermediate c. high
Response	<u>Response Demand</u> Overt-written Overt-spoken Covert	<u>Response-Demand Frequency</u> Infrequent-Frequent Frequency of response required a. low or zero b. intermediate c. high
Management	<u>Management Form</u> Repetition Multi-level Multi-form Error-diagnostic	<u>Management Frequency</u> Infrequent-Frequent Frequency of decision to change presentation a. low or zero b. intermediate c. high

[a] After Tosti and Ball, 1969.

c. Pictorial -- illustrative material, such as pictures

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.

Stimulus frequency (duration). Tosti and Ball (1969) 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.

Response demand. This dimension characterizes the types of behaviors which students are expected to perform in an instructional stiuation. 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.

Response demand 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 of 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 conceptualization of what has been termed "step size" than any other.

In a termporal 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 videotaped lecture.

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

Management form. Instructional management 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.

Every instructional system involves three management elements:

a. repertoire assessment -- appraisal of data and analysis of behavior compentencies

b. selection decision -- selection of a goal as a result of decisions based on assessment

c. activity -- actions following from decision.

It is evident that the elements of instructional management can vary in their composition, depending on the purpose of management. Tosti and Ball (1969) isolated five purposes that may be achieved.[1] 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

[1]Tosti and Ball (1969) originally identified a sixth type, that of operational management. In the current presentation, this category is included among the operational system variables, since the management activities are rarely contingent on assessment of an individual student's performance.

increase student learning rate.

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

The present research is primarily concerned with achievement management. Therefore, the four categories in the Management Type dimension presented below are the different procedures which may be used in 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. 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).

<u>Management frequency</u>. "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 more 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 any 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; decisions about a response cannot be made 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 Table 1)

Media Variables

Studies which have attempted to analyze differences between media have been largely conflicting and ambiguous. Reid and McLennan (1967), for example, reported 350 abstracts of media studies (mostly television and film); almost none of these studies found significant differences in media. Campeau (1967) selectively reviewed literature involving various comparisons among television, film, conventional lectures, programed instruction, pictoral presentations, radio and recordings, three dimensional models, and field trips. The large majority of studies reported no differences in student achievement and where differences were demonstrated (with the single exception of programed instruction) were as often in favor of one medium as another.

A number of researchers (Stolurow, 1962; Holland, 1965) have commented on the type of experimental comparison commonly attempted in studies of programed instruction. In most cases, these studies have attempted 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.)

This criticism applies to most experimental designs in which different media are compared. Media may differ in any number of ways, and be utilized in various ways. A programed instruction text, for example, presents relatively small units of material at a time, requires active responding by the student, may provide immediate feedback in terms of the correct answer, and may permit needed repetition of material. A film, on the other hand, is often viewed "passively" with large quantities of material presented in a short time, and rarely provides feedback or repetition. Even if there are differences in student achievement with these media, it is impossible to specify which elements of the instruction are responsible. In addition to the difficulty in interpreting demonstrated differences, the confounding of a large number of varied factors in "nonanalytic" comparisons also reduces the likelihood of finding any difference at all. The basis of this latter difficulty has been clearly explained by Campeau (1966):

> .....when a single medium is used to present an entire lesson, unit or course, and achievement resulting from essentially the same presentation by an alternative medium, it is quite feasible that each medium alternately succeeds and fails in supplying the unspecified array of learning events required for the various elements of the total learning task. Whether comparisons take into account effectiveness of media or methods, or identify special characteristics of learners and media which influence learning, it is furthermore quite feasible that over the duration of a lesson, unit, or course, the net result of these alternate successes and failures, when expressed as total criterion test scores, is to conceal real differences which do exist. Hence, perhaps the great preponderance of no-difference findings in media research.

The essence of Campeau's argument is that, when presentation variables are held constant, examination of media differences at the macro-level are unlikely to succeed, since the media differences which do exist in relation to particular learning tasks and students are opposite and counterbalancing. From this point of view, micro-analysis of media variables may succeed in demonstrating media differences at the level of the individual learning event.

Tosti and Ball (1969) take an even more radical view of media, based on the implicit postulate that there are no inherent advantages of media, but only disadvantages; i.e., that a medium only makes a difference when it places some limitation on the presentation design. From this point of view, the instructional systems designer should first establish a desirable presentation design, and then select media capable of delivering that instructional presentation. Given a constant presentation design, there should be no difference in student performance resulting from delivery of the presentation through different media even at a macro-level of analysis.

Briggs (1970) has developed a model for the design of instruction in which he places emphasis on the identification of the type of learning involved in each instructional objective. Analysis of the conditions necessary to bring about each type of learning aids in determining the media to be used. He argues that it is the responsibility of the educational specialist to define objectives and analyze learning types with sufficient precision to make obvious the necessity of particular media.

Briggs (1970) has presented a systematic method of working through the media-selection analysis, together with several examples of its application to a variety of objectives. Examination of these examples makes obvious the fact that selection at the micro-level involves a very difficult and detailed process, engaging the services of a highly sophisticated educational specialist for a lengthy period. Thus, as is the case for presentation design, design of optimal media-mixes for even a few simple behavioral objectives appears to be an expensive proposition when conducted at the micro-level.

It would seem very worthwhile, then, to see if, analysis at the macro-level could receive strong experimental support. Such results would serve to direct developmental efforts at the optimization of presentation design. If adequate levels of efficient learning can be achieved through presentation design at the macro-level of analysis, the need for expensive media-selection analysis and attendant expensive hardware may be bypassed, leaving future generations of educational technology to recoup relatively minor additional gains from refinement of media and presentation design at the micro-level.

In the present research, WLC has compared (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 in the second, the generality of conclusions such as Tosti and Ball's will be supported. Such findings would serve to redirect the general research effort in media; the question "Which presentation is more effective?" may be then considered more important than the question "Which medium is more effective?"

Task Variables

Basic to the development of the multimedia Leadership Management course was an explicit statement of educational intent or educational goals for students in that course. As Mager (1968) has pointed out:

> When clearly defined goals are lacking, it is impossible to evaluate a course on program efficiently, and there is no sound basis for selecting appropriate materials, content, or instructional methods. After all, the machinist does not select a tool until he knows what operation he intends to perform. ... Too often, however, one hears teachers arguing the relative merits of textbooks or other aids of the classroom versus the laboratory, without ever specifying just what goal the aid or method is to assist in achieving. I cannot emphasize too strongly the point that an instructor will function in a fog of his own making until he knows just what he wants his students to be able to do at the end of the instruction.

Mager defines "objective" as an intent communicated by a statement describing a proposed change in a learner--a statement of what the learner is to be like when he has successfully completed a learning experience. An objective is a description of the pattern of behavior, or performance, that the learner must demonstrate. Furthermore, a statement of the objective must denote the <u>measurable</u> attributes <u>observable</u> in the learner so two independent observers can infer correctly that the objective has been met. It is the observable and measurable character of instructional objectives which justifies the application of the term "behavioral" to such objectives.

A number of educational theorists have specified or implied that behavioral objectives involve different types of learning which may be arranged in a conceptual order from simple to complex.

Bloom (1956), for example, has written concerning his Taxonomy of Educational Objectives:

> 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.

In presenting his Taxonomy, Bloom distinguished two broad categories of objectives: (1) knowledge, i.e., the recall of specific information, and (2) intellectual abilities or skills, including comprehension, application, analysis, synthesis, and evaluation.

Along other lines, Gagne (1965) has developed a behavior 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

Gagne has argued that 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.).

In considering the effects of presentation design in relation to types of learning, it is important to carefully distinguish three major kinds of structure, or hierarchical organization involved in subject matter content and materials. (Briggs, 1968)
For convenience of discussion, these types of organization will be distinguished as involving content, products, and processes of learning.

The kind of organization involving content is the kind of logical arrangement of knowledge as might be conceived by an expert in the particular discipline. As Briggs has pointed out, the organization of knowledge as an outline of the field may be entirely different than the structure needed for learning purposes. A logical content outline is often a useful means of communication between professionals who "know about" the subject-matter of the field, but is meaningless as a guide to the novice, and has no necessary relationship to the types of learning required of the novice in gaining knowledge of the field.

The organization involved in the products of learning is more related to instructional design, and involves the interrelationships among behavioral objectives which are chosen for accomplishment by the student. Analysis of these compentencies to be achieved in behavioral terms is indicative of the kind of sequencing and arrangement of elements of instruction necessary to promote efficient learning and transfer of component competencies. Questions related to this kind of hierarchical structure concern what to teach and in what order.

The process kind of organization involves the nature and sequencing of the learning events required to attain the desired competencies, i.e., how to teach what is to be taught. Questions relevant to this kind of organization primarily involve the selection and arrangement of stimulus, response and management events designed into the materials to bring about processes resulting in a given learning product.

There is obviously an intimate relationship between the content, products, and processes of learning, but it should be clear that classifications such as Bloom's and Gagne's, refer only to the products of learning. Since classifications of products are of substantial value in the development of instruction as an aid in the analysis of content and design of materials, considerable confusion has arisen as to nature of the learning types which they identify.

It should also be pointed out that the content, products, and processes discussed above, are strictly speaking not those of learning at all, but are in fact the content, products, and processes of instruction, as conceived by an instructional designer. The structure resulting from the designer's analysis is represented in the materials developed to bring instruction to the learner, but the learner's actual behavior and modification in contact with the materials may be considerably different than that intended, even for very successful instructional materials. Since one of the goals of the behavioral analysis of instruction is to ultimately increase the correspondence between the structure of instructional components and those of the behavior of students, the present discussion will continue to refer to "types of learning." However, some confusion may be eliminated if it is kept clearly in mind that "types of learning" refer to types of instructional product as defined in stated instructional objectives.

In the light of the distinctions discussed above, a full discussion of task variables would require a systematic analysis of content-, product- and process-related variables, and of the relationships among them. The emphasis of the present research, however, was on the relationship between products and processes. The conditions of instruction investigated were designed to affect the processes of learning in a comprehensive fashion, altering the instruction related to every objective of a segment in specified ways. The main question raised by the classifications of types of learning is whether or not the presentation variables and media have similar effects on the achievement of different types of objectives.

In order to address this question it was necessary to develop objectives and related test items measuring achievement of different types of learning. On consideration of the large-scale manipulation of conditions of instruction, the development of tests representing rather large classes of types of learning was felt to be most in keeping with the general design of the research. The finding that particular presentation forms and media had different effects on very narrow classes of behavioral objectives would remove the advantage of instructional design at the macro-level; thus, it was desirable to determine if substantial effects of the conditions of instruction could be demonstrated using tests which include items representing several types of learning.

On the other hand, the finding of different effects with broad classes of objectives would support the procedure of segregating instruction on particular content according to the type of learning. Then the preparation of instructional units could proceed with large-scale control of presentation and media for each unit appropriate to the class of learning involved, and without major analysis at the micro-level.

Early in WLC analysis of content and objectives for the USNA Leadership Management course, it became apparent that most of the desired terminal objectives of the course could be placed at levels 7 and 8 of Gagne's (1965) hierarchy, with enabling objectives at levels 4 through 6. It was also recognized that most of the elements involved in these objectives could be further analyzed as rules and examples in accordance with Evans, Homme, and Glaser's (1962) RULEG system. Furthermore, the objectives could be further identified according to Bloom's (1956) taxonomy as involving either knowledge of leardership and management elements, or application of those elements in realistic naval situations.

As a result of these findings, WLC developed a system of formats to be used as guidelines in the writing of specific enabling and terminal objectives. This classification scheme represents a derivation and extension of Bloom's (1956) Taxonomy, Gagne's (1965) learning types, and Evans, Homme, and Glaser's (1962) RULEG system, serving to coordinate features of each.

Behavioral objectives were prepared, in most cases, in accordance with the four formats listed below:

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.

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.

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.

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.

Test items were developed from the statement of objectives for the criterion-referenced Progress Check tests and Administrative test which served as direct measures of achievement for purposes of evaluation. Performance on these tests thus represents learning of all four types given above, when the instructional materials are developed to explicitly teach those objectives.

The primary variables representing different classes of learning, however, were the Cumulative Posttest (CPT) items developed as special norm-referenced research tests. These tests were developed to have approximately equal numbers of items representing acquisition of knowledge (Type I item), and application of knowledge (Type II items), roughly corresponding to Bloom's categories of knowledge and applications. As items designed for content validity with high discriminative power, both types of items taps abilities in Bloom's other categories of comprehension, analysis, synthesis and evaluation.

Operationally, the distinction between Type I and II CPT items rests almost entirely on the presence or absence of naval situational examples in the stems or distractors. Thus there are some items which do not correspond precisely to Bloom's distinction between knowledge and applications.

The use of this operational criterion of distinction, however, appeared compelling after a careful analysis of content represented in the behavioral objectives. The special instructions and formats use in the preparation of CPT items are given in Appendix A.

The specific question raised in the present research is whether media and presentation variables have similar effects: (1) averaged over specific criterion behavior of all types as indexed by the Progress Checks, and (2) on criterion-related behavior involving comprehension, analysis, synthesis, and evaluation of knowledge and the application of knowledge as measured in the CPT tests.

Student Variables

The central idea motivating research into the relationship between student variables and instructional effectiveness has been to find methods of better tailoring educational systems to the needs and abilities of individual students. Obviously, this is an area of concern intimately related to the management of instruction, but the emphasis here is on determining what student characteristics can be assessed to permit management decisions, rather than on what decisions to make given some data on the student.

Several approaches to this problem have been reviewed by Cronbach (1967). Historically, there has been much interest in selection for advancement or ability-grouping, and for this reason, research largely centered around variables predicting general academic success. On the basis of such predictors, low-ability students have been weeded out, or assigned to courses of instruction of lesser difficulty or longer duration.

An alternative approach has been to assess individual long-range goals, and areas of ability and interest, and to provide optional courses of study which appear suitable for the individual. This has been the general approach of guidance and advisement programs, providing impetus for much research on tests in the areas of differential aptitudes and interests. More recently, this approach has been the basis of the development of large-scale computer-managed-instruction (CMI) systems, such as Plan (Brudner, 1969.) However, CMI systems are yet too new to assess their ultimate impact on individual-differences' research, since such systems have been operated primarily on the basis of a direct assessment of areas of competance, leaving the selection of goals to the teacher and student.

Only recently has major interest developed in a third approach involving the selection of a particular instructional method optimizing individual progress toward preselected goals. In the past, the selection of instructional method has been perogative of the teacher, who inevitablely modifies and utilizes methods according to his own abilities and history of success with various methods. Without standardized conditions, research on student variables predicting success under particular conditions has been difficult, if not impossible.

As Cronbach (1967) pointed out, individualized prescription of a method of instruction requires that alternative conditions of instruction designed for the same subject matter be compared in relation to student vairables to discover interactions between method and student. That is, one should seek to discover variables for which students in one score range find one condition superior, and other students in another score range find a different condition superior.

The recent developments in the use of standardized programmed instructional materials have provided the necessary context for meaningful reserach into student-method interactions. Findings in this area have been reviewed by Stolurow and Davis (1965) and Briggs (1968).

Sufficient evidence is available to conclude that student-method interactions are quite common, if not the rule. Interestingly, variables in the areas of personality, motivation, and attitudes appear to be as important, or more important than traditional academic predictors in the findings reported thus far.

In the context of the USNA Leadership Management course, the question of general academic performance is largely moot. The students at the USNA represent a select group in terms of academic ability, and it is unlikely that variables predicting academic performance would relate to any aspect of performance in the Leadership Management course.

The purposes of research on student variables in the present case concerned the prediction of overall course performance, and the prediction of achievement with particular media and presentation forms. Because of the number of conditions of instruction compared in the Leadership Management course, an invaluable opportunity was provided for one of the first large scale investigations of student-method interactions. To this end, a large battery of potentially predictive variables was included in the student data base. A detailed discussion of the design of the predictive battery will be included in Part II of this report, along with results of the correlational studies.

First, the investigation attempted to identify variables predicting final course achievement. Such variables may permit the identification of students unlikely to attain satisfactory levels of course performance. Further investigation of the source of difficulty for such students may be used to find some means of remedying their deficiency. The investigation of overall performance was of general educational interest, as well, since there are few previous studies of the prediction of course achievement in the area of the social and management sciences.

Second, student variables were related to performance with particular media. Such investigations provide information relevant to the assignment of alternate media, and on further investigation of students performing poorly with particular media, may also provide some suggestions for better accommodating particular media to the needs of individual students.

Finally, relationships between student variables and achievement with various presentation forms were investigated. The findings of these investigations may permit the utilization of the existing alternative presentations in an individually managed instructional system. In addition, some basic insights into the strengths and weaknesses of particular forms of instruction for individual students may be achieved.

Operational System Variables

A wide variety of variables must be considered when implementing an instructional system. In the traditional system the main variables dealt with are the scheduling of classes and the assignment of students and instructors to these classes. Many variables such as the length of the class periods, the grading system utilized, and the procedures for student interaction with the instructional materials are fairly well standardized.

The implementation of an individualized, multi-media instructional system necessitates a re-evaluation of many of these variables. The most obvious change involves the procedures for student interaction with the instructional materials. The degree to which students are allowed to work at their own pace must be determined. With students working at their own pace a logistical system for keeping track of the students as well as the materials must be established. Since the materials being used are not the typical "text," and since the mode of presentation used is typically not the lecture method, procedures must be established for guiding the student flow through segments of material where a variety of media are used. Once the full procedures for student interaction with the instructional materials and media are determined, consideration must be given to the personnel and facilities needed to implement those procedures. With respect to personnel, it must be determined how many students a single instructor can monitor and tutor and what additional personnel (if any) are needed to assist with record keeping and scheduling. It should also be recognized that the types of facilities as well as their arrangement will by necessity differ from those of the traditional classroom.

In a traditional classroom, time is typically held constant while performance is allowed to vary. In an individualized system, performance is held constant in the sense that all students must reach a predetermined level of performance, and the time a student spends or invests in reaching this level of performance is allowed to vary. This points out the possibility of utilizing a different set of variables to determine grades. One might, for example, base grades on the amount of time and number of attempts a student makes in achieving the desired level of performance. Final course achievement might also be based on the number of objectives achieved beyond the basic number required. If grades are indeed necessary, the type of evaluation system employed can serve as a very effective motivational device.

It should be noted here that the nature of the research involved in the course placed some artificial restriction on the operational systems variables. The actual operational system utilized will be discussed in sections two and three of this report.

III. RESEARCH PLAN

The objective of WLC's plan of research in the USNA Leadership Management course was to obtain experimental evidence relevant to the following general empirical questions:

a. Are substantial effects on student achievement produced by manipulation of presentational variables at the macrotaxonomic level as conceived by Tosti and Ball (1969)?

b. Are substantial differences in student achievement produced between different media delivering the same presentation, when measured over segments of material typical of a unit of instruction in most educational systems?

c. Are variations of conditions of instruction in the presentation design domain of greater or lesser importance than variation in the media domain?

d. Are the effects of presentation and media variables generalizable over different types of instructional objectives, or are different effects produced in relation to the acquisition and application of knowledge?

e. Are effects of particular presentation conditions and media similar for students varying according to established standarized tests of individual differences, or do the optimal conditions of instruction differ for different students?

Simultaneous accomplishment of research relevant to all of these objectives within a single ongoing course presented a number of difficulties requiring a complicated research plan. Several considerations important both to the achievement of clear-cut research findings and to the educational objectives of the USNA students in the Leadership Management course were taken into account in the development of WLC's research plan.

In performing several experiments within a single course sequence requiring repeated use of the same students it was necessary to arrange the experimental manipulation of materials and measurements so as to avoid the mutual entanglement of the effects of different experiments. Substantial variation of the level of difficulty in particular course content and test items required control to prevent obscuring of experimental effects. The small number of students available for enrollment in a developmental course required that special techniques for reducing random variation be employed to increase the precision of the experimental comparisons, yet without interfering with the investigation of individual differences in relation to experimental variables. Finally, experimental procedures were needed which would not place an excessive burden of time and effort on the individual student, nor handicap his overall achievement through placement in ineffective learning conditions, thus leading to an undeserved reduction in course grade.

On careful consideration of all factors, a research plan was devised which substantially satisfied the criteria given above with minimal compromise among objectives. The ability of the research plan to reconcile such apparently contradictory requirements commends the WLC design approach as a model for research in ongoing courses undertaken under similar limitations.

In the outline of the research plan below, the structure of the course is described in the first subsection, with particular attention given to the cumulative posttest (CPT) unit which served as the basic research unit of instruction. The next subsection describes the general approach taken in superimposing controlled experimental procedures on the course structure. This is followed by a description of the procedures used in statistical evaluation of experimental results, including methods of handling missing data, and loss of balance resulting from dropped students. The final subsection outlines the approach used in evaluating the student characteristics in relation to overall course achievement and performance under particular instructional conditions.

Course Structure

The Leadership Management course was first organized in terms of elemental blocks of content and related tests of student achievement, which were temporally sequenced without regard to research constraints. Additional elements of structure were then inserted for research purposes. This procedure insured that a basic course structure was achieved from which the research elements could easily be detached for purposes of final course packaging and implementation. The course structure may be described in terms of the four categories outlined below.

Part. The content is divided into 12 parts, corresponding to 12 chapters of the basic content outline. Each part is a formal designation of a large topic area, representing a substantial number of closely related terminal objectives relatively independent of the objectives of other parts. The objectives of any one part could be considered to be subsumed under one of the broad aims (macro-objectives) of the course. The part served primarily as an aid in fractionating the developmental work on materials.

Segment. In terms of content, a segment is a sub-collection of learning objectives within a part, which are closely related in the development of a behavioral hierarchy of competence and in the sequencing of instructional events. A total of 59 segments were incorporated in the 12 parts of the course. The content headings of each segment are listed in Table 2 under their respective parts.

Conceived operationally, the segment is the basic instructional unit in the development and production of materials, and serves as the logistical unit in implementation for purposes of scheduling and assessment of progress through the course materials. Essentially, the segment is analogous to a class period or lesson in other instructional systems, requiring 40 to 80 minutes of student time, and provides the basis for manipulation of the real-time parameters of the course.

At the completion of each segment, a progress check (PC) test is administered to assess the student's attainment of the terminal and enabling objectives of the segment. PC's are composed of 10 criterion-referenced items, developed directly from the behavioral statement of segment objectives.

Table 1

OUTLINE OF COURSE STRUCTURE AND MEDIA

Part and Segment Number	Content Heading	CPT Unit[a]	Medium[b]
	PART ONE: OVERVIEW OF LEADERSHIP		
1.1	Concepts of Leadership	NR	ST
1.2	Standards of Leadership in the Naval Service	NR	F-GD
	PART TWO: INDIVIDUAL BEHAVIOR		
2.1	Introduction to Psychology	NR	ST
2.2	Behavior and Its Observation	1	AT- or VT-PB
2.3	Learning	1	AT- or VT-PB
2.4	Factors Affecting Learning	1	AT- or VT-PB
2.5	Attention and Perception	1	AT- or VT-PB
2.6	Motivation	2	ST
2.7	Conflict	2	ST
2.8	Neurotic and Psychotic Reactions	2	ST
2.9	Personality	NR	LAS
	PART THREE: GROUP DYNAMICS		
3.1	Characteristics of Groups	3	AT- or VT-PB
3.2	The Relationship of the Leader to the Group	3	AT- or VT-PB
3.3	Group Interactions	3	AT- or VT-PB
3.4	Conformity as a Factor of Group Behavior	3	AT- or VT-PB
3.5	Relation of the Individual to the Group	NR	ST
	PART FOUR: ACHIEVING EFFECTIVE COMMUNICATION		
4.1	Importance of Interpersonal Communication	4	LT
4.2	Types of Communication	4	LT
4.3	The Communication Process (Receiver and Barriers)	4	LT
4.4	The Communication Process (Sender and Feedback)	5	AT-IPB
4.5	Formal Communication and Its Dimensions	5	AT-IPB
4.6	Informal Communication	5	AT-IPB
4.7	Communication Under Battle Situations	5	AT-IPB
	PART FIVE: MILITARY MANAGEMENT		
5.1	Introduction to Management and the Management Process	NR	ST
5.2	Decision Making and Creativity	NR	ST
5.3	Objectives	NR	ST
5.4	Planning	6	LT
5.5	Organizing: Principles and Process	6	LT
5.6	Organizing: Structure	6	LT
5.7	Organizing: Charting	7	AT- or VT-PB

Part and Segment Number	Content Heading	CPT Unit[a]	Medium
	PART FIVE: MILITARY MANAGEMENT (CON'T)		
5.8	Directing	7	AT- or VT-PB
5.9	Controlling	7	AT- or VT-PB
5.10	Coordinating	7	AT- or VT-PB
	PART SIX: AUTHORITY AND RESPONSIBILITY		
6.1	Concept of Authority	8	ST
6.2	Why People Accept/Resist Authority	8	ST
6.3	Delegation of Authority; Line-Staff Relationship	8	ST
6.4	Responsibility	NR	ST
	PART SEVEN: LEADERSHIP BEHAVIOR AND STYLE		
7.1	Leadership Behavior	9	AT- or VT-PB
7.2	Leadership Style	9	AT- or VT-PB
7.3	Determiners of Leadership Style - The Leader	9	AT- or VT-PB
7.4	Determiners of Leadership Style - The Group and The Situation	9	AT- or VT-PB
7.5	Participative Leadership	NR	VT-PB
	PART EIGHT: SENIOR-SUBORDINATE RELATIONSHIPS		
8.1	Organizational Structure & Social Distance in Senior-Subordinate Relationships	10	LT
8.2	Officer-Enlisted Relationships	10	LT
8.3	Assumption of Command and Formal & Informal Leader Relationships	10	LT
8.4	Introduction to Counseling	11	LAS
8.5	The Counseling Process	11	LAS
8.6	Relations with Seniors and Contemporaries	11	LAS
	PART NINE: MORALE - ESPRIT DE CORPS		
9.1	Morale	NR	VT-PB
9.2	Group Solidarity and Esprit	NR	VT-PB
	PART TEN: DISCIPLINE		
10.1	Introduction to Discipline	NR	AT-IP
10.2	Development and Maintenance of Discipline	NR	AT-IP
	PART ELEVEN: PERSONNEL EVALUATION		
11.1	The Role of Evaluation	12	ST
11.2	Enlisted Performance Evaluation	12	ST
11.3	Officer Evaluation	12	ST

Part and Segment Number	Content Heading	CPT Unit[a]	Medium
	PART TWELVE: APPLIED LEADERSHIP		
12.1	Measurement of Effective Leadership	13	CAI
12.2	Generally Recognized Characteristics of an Effective Leader	13	CAI
12.3	Techniques of Assuming Command	13	CAI
12.4	"That's an Order!"	13	CAI

[a] NR refers to a nonresearch segment, thus not assigned to a CPT unit.

[b] ST=Syndactic (multi-level) Text; F-GD=Film, Group Discussion; AT=Audiotape; VT=Videotape; PB=Panelbook; LAS=Learning Activities Summary; LT=Linear Text; IP=Intrinsic Program; CAI=Computer Assisted Instruction.

<u>Module</u>. A module is a particular instructional condition used to prepare and deliver materials for a segment, identified in terms of the categories of the Tosti and Ball (1969) model. Several parallel modules were prepared in each segment utilized for research purposes, representing variations specified by the experimental designs. The different modules of a segment are distinguishable from one another by differences in presentation design and/or media, although the content is the same. Specifications of the modules for each segment are outlined in later sections of the paper giving the design of each experiment.

<u>Cumulative posttest unit</u>. The cumulative posttest (CPT) unit is a group of three or four adjacent segments within a part. There are 13 CPT units involving 45 of the 59 segments of the course, as listed in Table 2. The primary criteria for grouping segments into CPT units were that the segments dealt with similar types of content and objectives, and that the instructional sequences relating to particular concepts which were initiated in the unit would also terminate in the same unit. All segments in a CPT unit were developed in the same medium and with the same variations in instructional conditions between modules.

The CPT unit is the fundamental unit of instruction for research purposes, providing the framework on which the experimental designs were constructed. The students were divided into groups assigned to different modules in the CPT unit. A student in any one group would thus encounter the same experimental conditions in progressing through the three segments of the unit, and would take three PC's, one after completing his module of each segment. After completing the segments and PC's all students then take the CPT, a test administered to assess overall achievement level under the experimental conditions represented in the CPT unit.

Performance on the CPT was the primary dependent measure for research purposes. Each CPT was composed of 10 multiple-choice items for each segment in the unit, so that CPT's for 3 segment units had 30 items, and CPT's for 4 segment units had 40 items. There were approximately equal numbers of two types of items: Type I, representing acquisition of knowledge of the concepts and principles in the unit, and Type II, representing application of those concepts and principles in the unit in relation to realistic examples of leadership situations.

CPT items were designed to have content validity in relation to the objectives of the unit, but unlike the PC items, also to have high difficulty and discrimination power. The CPT tests thus provided norm-referenced rather than criterion-referenced measures of achievement level. Many items were designed to measure the ability to integrate behaviors from different segments in the unit. An effort was made, however, to maintain an equitable representation of content from the several segments of the unit.

Following completion of the CPT, each student is given remediation on segments where his PC test performance is below 80%. The remediation consists of repetition of the same instructional materials previously used with the segment, or materials of an alternative module thought to be more effective. On completing remediation, the student repeats the PC's for those segments and then proceeds to the next segment.

Media. Course materials were prepared using eight media:

a. videotaped lectures with accompanying panelbooklet (VT-PB)

b. audiotaped lectures with accompanying panelbooklet (AT-PB)

c. syndactic text -- summary statement on concepts in text followed by frame sequence in text (ST)

d. audiotape with intrinsically programed text (AT-IP)

e. computer assisted instruction (CAI)

f. linear programed text (LT)

g. learning activities summary -- annotated bibliography and resource materials (LAS)

h. film and group discussion (F-GD)

The nature of these media, along with other media in nonresearch segments are explained in more detail in the Phase II Evaluation Report (TR-6.11) on the U. S. Naval Academy Leadership Management course.

After media selection and analysis of course content, and the determination of the number and size of CPT units needed for the research design, CPT units were designated within each part, as listed in Table 2 . For example, Part II has nine segments, and it was possible to arrange one four-segment CPT unit and one three segment CPT unit in Part II, leaving two segments as nonresearch segments. Media appropriate to unit content and providing a varied sequence of media were then assigned to the units.

As may be seen in Table 2, the same medium was used with all segments in a CPT unit, and except for two cases (Segments 6.4 and 7.5), the medium changed at the completion of a CPT unit. Ordinarily, at least two other media were used before the presentation returned to a given medium. Thus, there was considerable media variety in each part of the course, and a fairly balanced distribution of media placement.

It may also be noted that while students might experience different sequences of modules in moving through the CPT units, all students received the same media sequence with the single exception of the parallel AT- and VT-PB media.

Experimental Design.

The experimental designs used to arrange the experimental instructional conditions in the research plan involved several common principles which were followed insofar as possible.

a. Conditions of instruction of primary experimental interest were always compared between alternate modules in the same CPT unit.

b. Students were randomly arranged in groups assigned to alternate modules in the CPT units. The primary test of achievement was the CPT, which measured achievement over the entire unit under the conditions of a constant module.

c. Several widely separated CPT units involving the same medium were used in each experiment, with the same modules appearing in each unit. Experiment V, which involved only one CPT unit was an exception to this rule.

d. Over the CPT units of a given experiment, each group of students experienced all types of modules involving the conditions of primary interest in one sequence of a counterbalanced set of sequences. Experiments III and V were exceptions to this rule.

In technical terms, these principles may be summarized by the statement that Experiment I, II, and IV were designed as various types of repeated measure Latin Squares with CPT units defining the columns of the squares, randomized groups assigned to the rows of the squares, and modules corresponding to the counterbalanced latin letters of the squares. Experiment III was a mixed repeated-measure randomized block design, with repetition of modules and CPT units as blocks, and Experiment V was a completely randomized design (Meyers, 1967).

It is the purpose of this section to outline the basic problems of experimental control and course administration which lay behind the decisions to design the experimental conditions and sequences of events according to these rules. It is believed a detailed consideration of the factors involved and the methods applied in the present case may provide considerable guidance in the design of research conducted in on-going course systems.

Design of the CPT unit. The CPT unit was so arranged as to accomplish several objectives relating to the effectiveness of the research and to course administration.

Two considerations were involved in the decision to use a group of segments as the basic research unit, rather than single segments. First, the unit of several successive segments more closely simulates the procedure under which the results of the research would be applied, where conditions found to be superior would be held constant over large sections, if not a whole course. Thus, it was desirable to provide conditions permitting the detection of delayed or slowly developing effects which might not appear immediately on the first administration of the conditions. Second, the single PC test was unlikely to provide a measuring instrument of sufficient sensitivity to demonstrate real efforts of the varied conditions of instruction. Limitations on the total student time which could be devoted to testing required that each PC be short, and given that effective materials were prepared, there should be relatively little variation in immediate achievement on items directly referenced to segment objectives. Relatively low test reliabilities could therefore be expected both on the basis of test length and the criterion-referenced nature of the items.

The CPT was constructed to provide a norm-referenced test of greater reliability and discriminative power than could be achieved even by aggregating PC items across segments. Testing achievement of the unit as a whole following completion of the unit not only permitted testing of interrelationships and integration of the content of the segments at a higher level of complexity and difficulty, but provided measures at a point where some retention loss of learning in the early segments would have an opportunity to occur. The placement of the CPT at the end of a series of segments, then, could be expected to assess more accurately the amount of retainable learning achieved than would any testing conducted at the end of each segment. The contribution of retention loss to CPT performance would also be expected to increase sensitivity of the test to effects of the experimental conditions.

The decision to develop parallel modules within segments and to hold the medium constant over the unit was predicated on basic considerations of experimental control. Segments of course content could be expected to vary considerably in difficulty and performance level. Additional variation would be contributed by variation in the effectiveness of materials developed to teach that content, and in the difficulty level of test items measuring achievement. The use of parallel modules with common content meant that conditions could be compared which varied only according to the presentation variables intended, with all other details of the presentation held precisely equivalent. In this fashion, most variation from sources associated with content was removed from the experimental comparisons into the columns of the Latin Squares, thereby enabling much smaller differences between conditions to be declared as statistically reliable.

Although the presentation variables can be regarded as fundamentally similar in different media, they cannot always be implemented according to exactly equivalent rules or criteria. The use of a single media in the CPT unit meant that the presentation variables could be manipulated in precisely the same way in each segment of the unit. Variation in the meaning and effects of the presentation variables was thus avoided, producing both a further gain in statistical precision, and a more explicit and unambiguous realization of the categories of the presentation variable, allowing easier interpretation of results.

It was, of course, necessary for groups of students to be assigned to a single module throughout the CPT unit so that their CPT performance would reflect the influence of only one constant set of experimental conditions. Since the student also changed media whenever beginning a unit, special instructions relating to the module could be conveniently incorporated in special instructions relating to the media without drawing undue attention to the variables being manipulated.

An additional consideration determining the plan of the CPT unit concerned the role of remediation. Remediation was required to bring all students to similar levels of competence following the measurement of experimental effects on the CPT. This procedure insured that no student would be disadvantaged by assignment to an inferior module except through loss of additional time for remedial learning. Furthermore, remediation insured that substantial differences in achievement among modules would not be carried over to the next research unit to confound the differences between modules compared there, nor to increase variation among students thus reducing the precision of the statistical evaluation.

The final point to be noted is the relation of the CPT unit to the final course system. The major features which demark the CPT unit are parallel modules and the CPT test. Since a single superior module may be selected and the alternatives abandoned, and the CPT may be eliminated, it should be clear that the CPT unit is readily adaptable to dismemberment in the ultimate use of the developed materials. The only trace of the CPT unit then remaining would be the points at which media change.

Media selection and sequencing. Selection of media for the instructional design of the USNA Leadership Management course was predicated on the requirements of the experimental designs, capacities needed for delivery of instructional presentations, and the diversity and flexibility expected of an individualized multimedia course. The media selected permitted precise experimental control and planned variaton in dimensions of stimulus representation, duration, response form, response demand frequency, and management decisions. Within limitations of existing facilities at USNA, media were selected which can be used in individually paced instruction without undue logistic difficulties, and with sufficient variety of instructional technique to maintain a consistent level of student motivation. Some media were selected for their novel appeal, while the experimental manipulations of presentation design provided variety in the utilization of more traditional media forms.

Media placement and sequencing was limited to some extent by the number of segments required for each CPT unit and the number of segments in each part. Within these limitations, media were assigned to ensure perception of a sense of media variety, and to provide persistent media in segments with the most complex concepts.

The final media assignment to segments was carried out so as to provide a sufficient number of CPT units in the same media to accommodate the designed experiments, and to keep the CPT units of a given experiment widely separated in the course.

The purpose of having widely separated units in the same experiment with units assigned to other experiments intervening was to insure that any carryover effects of experimental conditions which remained after remediation were not carried systematically into the treatment conditions of the same experiment. Such carryover effects were, instead, randomized among the treatment conditions of different experiments. Thus, while the carryover effects might produce some increase in variability and loss of precision in comparison of conditions of instruction, they were not allowed to systematically bias any comparison of conditions. Wide separation of CPT units in the same experiment also could be expected to systematically reduce problems of sequential correlation often associated with experiments of the repeated measure type.

The purpose of assigning the same medium to the CPT units of a particular experiment was the same as that given for holding the medium constant in the CPT unit, i.e., to permit the presentation variables to be manipulated in the same way and to produce similar effects in each segment of a given experiment. Since the presentation variables might possible have different effects in different media (media-presentation interactions), such variation in effects was avoided in all but Experiment III, where the interactions were of direct interest. Otherwise, the linear models for the statistical analyses would have been based on erroneous assumptions, and the interpretation of results would be somewhat difficult.

Arrangement of systems and modules. Several advantages accompanied the use of counter-balanced sequences of modules across the maj~rity of CPT units. First, each student encountered all the major module variations, roughly equating exposure to relatively good or poor conditions of instructions. In addition to remediation, the equation of experimental histories produced by counterbalancing ensured that no student was handicapped in opportunity to obtain good grades through consistent assignment to inferior conditions. This was a distinct administrative advantage in the assignment of grades, since no special correction or subdivision of students was required to account for differential effects of experimental history.

Measurement of performance of each student on each type of module permitted the evaluation of the primary conditions of interest on a within-student basis. Within-student comparisons involve a marked gain in precision (reduction in random variability) since the variation among students does not contribute to differences between experimental conditions. The gain in precision produced by this means was especially important because of the small number of students available.

Experience with within-student designs also indicates that the treatment differences found in such designs tend to be more characteristic of all individual students and less fragile in the face of alterations in the context and preceding events. Only relatively consistent and durable experimental differences remain after averaging over such diversity of content, and module and media histories as are produced by the counterbalanced variations.

With the gain in precision produced by counterbalancing, further gains from matching groups or statistical control through analysis of covariance were not deemed necessary, permitting random assignment of students to groups. There was no assurance that any variable available for matching or analysis of covariance would have a sufficient relationship to CPT performance to produce any substantial gain in precision. Furthermore, the use of either of these procedures would have increased the complexity of the design and/or statistical analysis to the point of unmanageability. Corrections for student withdrawals and missing data, as discussed below, would have been much more difficult as well.

Two other considerations also favored the use of randomized groups. Evaluation (with lesser precision) of certain media and presentation main effects and interactions of secondary interest was permitted on a between-student basis through judicious arrangement of the Latin squares. Furthermore, randomization provided that carryover effects, as discussed previously, were not allowed to systematically enter the comparison of conditions of instruction.

Analysis of Variance Methods

Four sets of data on each experiment were analyzed, including PC data totaled over the segments of each CPT unit, total scores, CPT Type I scores, and CPT Type II scores. Wherever the test scores were based on differing number of test items for separate CPT units, the original scores were concerted to percentages prior to analysis.

Analysis of variance was performed on each set of data based on standard linear model methods for the types of designs involved (see especially Winer, 1962, Ch. 7, 8, and 10, and Meyers, 1966, Ch. 8, 9, and 10), with some modifications required as described below. Although multivariate analysis was jointly applicable to the different measures obtained, the univariate analyses were preferred for ease of computation and interpretation in the light of the complexity of the designs.

Inspections of residuals indicated reasonable satisfaction of the required statistical assumptions, so no statistical tests of these assumptions were performed.

The small number of conditions compared in any one experiment also obviated the need for multiple-comparison procedures to aid in the interpretation of results found significant by overall F-tests.

Two problems did arise, however, which required special techniques of analysis. First, there were several cases in which individual CPT scores were inaccurate resulting from minor errors in implementation of experimental procedures. Such scores were dropped from the analysis and replaced by least squares estimates. The estimates were obtained by following the procedure developed by Yates for the estimation of missing data in randomized block designs as given in Cochran and Cox (1957, p. 110). This procedure was appropriate since the data of a single group within a repeated-measure Latin square design forms a randomized-block design when students are identified as blocks. Computation was based on the two-way student x unit table from which the score was dropped.

The second problem was that four of the original 48 students withdrew from the course, resulting in unequal group sizes in all experiments. Since the loss of students was unrelated to the nature of the experimental conditions, and the group sizes which remained were not very disparate, the computational procedure for unweighted-means analysis of variance was followed in all cases (Winer, 1962, p. 374-378). In this form of analysis, components of error variance are estimated from the original data of each group, but the analysis of treatment effects is based on tables of unweighted means. The use of unweighted means causes all experimental conditons to contribute equally to the estimation of effects, without regard to the number of individuals in those conditions.

Analysis of Student Characteristics

Multiple regression analyses were performed to relate individual differences among students to learning achievement. Three types of dependent measures were used as the basis of these analyses.

First was the administrative posttest used as the criterion variable for prediction of overall course achievement. The second type of variable was the student total residual derived from average student performance in each experiment, and used as the criterion variable in prediction of achievement with a particular medium. The third type of variable was the within-student residual derived from scores on a module, used as the criterion variable in predicting achievement in a particular presentation design. The latter two types of variables are identified as sources of error variance in the analyses of variance and represent unexplained individual differences in student performance after overall module and CPT unit differences are removed. In every experiment, residuals were derived for total CPT scores, CPT Type I scores, and CPT Type II scores.

A total residual was obtained from a student's mean performance over all CPT units of an experiment by subtracting the mean of the group (to which the student belongs in that experiment) from the student's mean. The resulting deviation score represents how well the student learned in relation to his group over the entire experiment. Since each experiment involved a particular medium, this score indicates how well the student learns in connection with that medium, at least for the kinds of content and presentations used with the medium. Regression of the total residuals on the battery of student variables could thus be used to identify variables associated with variation in achievement with particular media.

A within-student residual was derived by subtracting the mean for the student's group in a particular module from the student's score in the module, and secondly, subtracting the total residual for the student from the result of the first subtraction. The resulting deviation score represents how well the student learned in relation to his average standing in the group, and in relation to the average performance of the group on that particular module. When the within-student residuals for a particular module are regressed on the battery of student variables, variables are identified predicting performance in the presentation conditions defining the module.

A battery of 138 predictor variables was used in the regression analyses. Included in the battery were common standardized tests in the major areas of aptitude, achievement, personality, motivation, and interest. Also included were items of student questionnaire data concerning two previous studies in related subject matter, experience with methods of instruction, and self-reports on study habits and academic abilities. Emphasis in the selection of tests was on commonly used and well-standardized tests, with considerable established validity to aid in the interpretation of findings. Emphasis in the selection questionnaire items was on face validity.

Because of the large number of predictors and the small number of students available, and the fact that little confidence could be placed in most *a priori* hypotheses relating performance and predictors the analyses of Phase II were designed as a variable selection process. The aim was to filter out potential important predictors from the many candidates available, thus setting the stage for a cross-validation of results in Phase III.

The analyses for each criterion variable were conducted in three stages. First, potential predictors and suppressor variables were selected on the basis of first-order correlations with the criterion and inter-correlations among the predictors. Second, the selected variables, plus fifteen others selected on an a priori basis were submitted to a step-up regression analysis in which variables contributing most to prediction are added one-by-one into regression until no further significant improvement in prediction is achieved. Variables surviving the step-up, plus the 15 a priori variables were then submitted to a step-down analysis, in which variables are eliminated one-by-one, according to which variable contributes least to prediction. Procedures of the step-up and step-down analyses are based on those described in Draper and Smith (1966).

The predictive battery, procedures of analysis, and results of regression analyses will be described in more detail in Part II of this report.

IV. RESEARCH IMPLEMENTATION

Students

Forty-eight third classmen (sophomores) from the United States Naval Academy were enrolled in the Leadership, Psychology and Management course. Four of these students dropped the course within two weeks of the start of the semester.

Before commencing work on the course materials, each student was randomly assigned to a track. This student track indicated the precise module of materials a student would use in each of the 59 segments of the course.

Materials

The basic course structure was discussed in Section III, and Table 2 presented an outline of the course structure and the media used. The segments listed in Table 2 are core segments. That is, they are required segments which include all of the information pertinent to the attainment of the requisite behavioral objectives. In addition to core segments there were depth core and enrichment segments. Depth core segments were associated with one or more segments and were directed toward amplifying the learning objectives of those segments. Depth core segments included in the first implementation were film, group discussions, and classroom lectures by the USNA instructor. Unlike core segments, depth core were scheduled by the instructor with respect to time and place. Student attendance was required. Enrichment segments were related to but not essential to the mastery of terminal objectives. They were optional to students who desired more information than that presented in core segments.

The specific media used for each segment are given in Table 2. A description of the media and each of the variations in presentation form (modules) within each medium is given in Sections V through IX.

Tests

Four different tests were used throughout the course. They were the administrative pretest and posttest, the progress check, the cumulative posttest (CPT), and the USNA examination.

The administrative pre and posttest was an 80 point criterion referenced test composed of items representatively sampled from the objective-test item pool. There was at least one administrative test item for each segment of the course. The pretest was given at the beginning of the course, and the posttest was given as part of the final examination.

The progress check was a criterion referenced test of approximately ten items. It was given at the end of each segment.

The cumulative posttest (CPT) was a norm referenced research test composed of positively discriminating content-related test items. Each CPT was composed of ten items for each segment in the unit. Cumulative posttests were given at the end of each experimental unit.

USNA examinations were a combination of criterion referenced test items selected from the objective-test item pool and items developed by the USNA on-site instructor. These were the only tests in the course which were used to determine the midshipmen's grades.

Procedures

The first implementation of the course was conducted in the Spring of 1970. The course was administered by the USNA on-site instructor, the WLC on-site instructor, the WLC systems analyst, and two data clerks. Detailed procedures used in implementing the course are given in the Instructor's Guide (TR-6.6).

The instructor's basic responsibilities were tutoring students needing remediation, leading group discussions, scheduling and administering depth core segments, scheduling and administering examinations, and determining grades.

The systems analyst developed and supervised the logistical procedures of the course. Data clerks controlled dissemination and collection of all core materials, remediation prescription forms, module questionnaires, progress checks, and cumulative posttests (CPTs). They also scored progress checks and CPTs and forwarded data to WLC's computer center.

Students were routed through the course according to procedures outlined in the Student Guide (TR-6.5). In brief, students worked through core segments of the course at their own speed. They were allowed to check out software materials and study them whenever and wherever they wished. All students were given identical material when they studied a non-research segment; i.e., they were instructed by the same form of presentation. For research segments, they studied by the particular module (form of presentation) to which they were assigned. Students were randomly assigned to modules at the beginning of the course. Each student received his own routing schedule which included not only the sequence of segments he must study but also the schedule for remediation, research tests, and USNA examinations.

Students worked through non-research material by studying a segment, taking a progress check, remediating (if necessary), and then retaking the progress check. The requirement for remediation was based on failure to attain 80% of the objectives as measured by the progress check. If the student failed to meet the 80% criterion on his first try, he was given a remediation prescription form which directed him to specific points in the materials which related to the objectives failed. If the student failed to meet the 80% criterion following remediation, he reported to the on-site instructor for tutoring.

Students worked through research segments in the same manner as non-research segments except that they did not remediate until after they had completed the entire research unit and taken the cumulative posttest. Specific procedures to be followed in the research segments are discussed in sections V through IX.

Facilities

For the implementation of the course, WLC was provided three classrooms at the Naval Academy. One room which was designated as the administrative office contained desks for the administrative staff and storage space for half of the course materials (including tapes, printed material, tests, forms, and computer cards). The administrative room was used as the site for administrative conferences, for student tutoring, and for distribution and collection of all material.

The second room was used as the principal instruction room. It contained 15 student carrels equipped with Ampex VTR's (4900), TV monitors and earphones, and Craig cassette recorders.

The third room, used as a multi-purpose room, had three carrels to handle overflow from the instructional room. In addition, there were 30 student writing desks which were used during depth core lectures, films, group discussions, and testing.

V. EXPERIMENT I

The purpose of Experiment I was to evaluate the effects of variation in response demand frequency (RDF) on student achievement with transient media. Additionally, two transient media, videotape and audiotape lecture (each combined with a panelbook), were used to deliver the designed presentations, permitting direct comparisons between media under constant conditions of instruction.

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 primary hypothesis tested in this experiment was whether transient presentations with high RDF will be more effective than transient presentations with low RDF. 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 occurrance of the question per sé. 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 experiments in persistant media, as reviewed by Anderson (1967) and Gagne and Rohwer (1969), have established that the insertion of questions among segments of text can produce both general and specific facilitory effects on learning. The primary mechanism involved in such effects is thought to be attentional; the questions apparently serve to create or maintain classes of inspection behavior directed toward general or specific features of the material. One basis of this conclusion was the finding that the effects occurred even if knowledge of the correct answers was not supplied, although feedback did appear to increase the specific effects of questioning.

Studies in transient media also have demonstrated improved performance resulting from increased RDF and feedback. 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):

> ...anITV 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) 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 whether the test items were unrelated to the participation questions. 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.

In view of the fact that both the Gropper and Lumsdaine (1961) and Abbey, et. al. (1963) combined questioning and feedback, it remains to be determined if high RDF alone will produce facilitating effects in transient media similar to those found with persistent media. 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." Whether resulting from attentional effects or step size, the finding that high RDF is more effective than low RDF in transient presentation designs would support the utility of Tosti and Ball's (1969) taxonomy, and would provide a general guideline for the preparation of materials in transient media.

The transient media of videotape and audiotape lecture were selected in the present experiment primarily to provide an appropriate context for the test of the RDF hypothesis discussed above. However, a second important purpose of this experiment was to compare media delivering essentially identical presentation conditions. It therefore became necessary to incorporate a panelbook medium forming a media-mix which would compensate for the visual limitations of the audiotape medium. While the media compared are not "pure" in one sense, they do provide control of the conditions of instruction, and the results can be more readily applied in contemporary instructional procedures where mixing of stimulus forms and media is often used to compensate for the limitations of one media. As Tosti and Ball (1969) point out:

> ...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.

Specifically, the present experiment tested the hypothesis that there is no difference in student achievement when identical presentations are delivered via audiotape lecture with panelbook, or videotape lecture with panelbook. It may be noted that, for purposes of control of presentation form, materials in both media were prepared with zero management frequency and therefore no management types, as is most typical of videotape and audiotape media applications.

In each segment of material, the audiotape lecture delivered the audioscript of the videotape lecture. The panelbook in the audiotape condition was used to display charts, diagrams, or other visual aids essential to instruction as indicated by the content material, while these visual materials were presented on the students' viewing screen in the videotape condition. The panelbooks for both conditions also contained the questions representing high or low RDF conditions, as the case might be, and required overt-selected responses. Thus the presentations were constant in form and frequency of stimulus and managements, and form of response, but with both high and low RDF conditions.

A final question might be raised as to whether the selection of transient media does or does not impose some inherent limitation on effectiveness in those segments involved in Experiment I. Similar questions may be asked in relation to the other conditions held constant in the experiment. The ultimate answer to such questions of course, can only be gained by empirical investigation; nevertheless they are worth examining in the light of available evidence.

Tosti and Ball (1969) 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 pr sentation that lasts as long as the student requires, i.e., a student-paced presentation.

Following this rule, the transient presentations of Experiment I were carefully placed in segments of introductory or other relatively familiar content. Similar careful consideration was given to the effect other factors held constant in Experiment I to avoid other potential limitations on effectiveness.

Schramm (1964) has provided a brief review of some findings relevant to the general effectiveness of transient presentations. 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 per cent below and 10 per cent 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 well 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.

METHOD

Design and Analysis

The plan of the experiment is presented in Table 3. The students were randomly assigned to four groups, each of which received a particular combination of RDF and sequence of video- and audio-tape presentation in the four CPT units used in the experiment. Sequence 1 involved videotape in the first half of the CPT units (Units 1 and 3), followed by audiotape in the second half (Units 7 and 9), with Group 1 given the high RDF presentation and Group 2 the low RDF presentation. Sequence 2 presented audiotape followed by videotape, with Group 3 having high RDF and Group 4 having low RDF presentations, respectively.

Sequences and RDF are orthogonal between - student variables in this design. Unit halves and taped media are repeated measure variables arranged as a balanced crossover design. The crossover characteristic results in the confounding of certain interactions with the main effects and interactions of interest as shown in Table 4, which presents the breakout of sources and degrees of freedom for the analysis of variance. Since only eight out of the possible sixteen cominations of RDF, sequence, media and halves are employed in the experiment, the pattern

Table 3

Experiment I

Revised Plan of Experimental Design

Group	Response Demand Frequency	Cumulative Post Test Units			
		1	3	7	9
		Sequence 1			
1	HRDF	Videotape	Videotape	Audiotape	Audiotape
2	LRDF	Videotape	Videotape	Audiotape	Audiotape
		Sequence 2			
3	HRDF	Audiotape	Audiotape	Videotape	Videotape
4	LRDF	Audiotape	Audiotape	Videotape	Videotape

Table 4

Experiment I

Confounded Sources of Variance

in the Revised Model [a]

Source of Variance	Confounded Source (Alias)	Degrees of Freedom
Sequence (Q)	MH	1
Frequency (F)	FMQH	1
QF	MHF	1
Students (S/QF)	-	40
Total Between Students		43
Half (H)	MQ	1
Units (U/H)	-	2
Media (M)	QH	1
QU/H	-	2
FH	FMQ	1
FU/H	-	2
MF	QFH	1
QFU/H	-	2
HS/QF	-	40
US/HQF	-	80
Total Within Students		132

[a] Based on pattern of confounding of a one-half fractional replication of a 2^4 experiment derived from the defining contrast MQH.

of confounding may be analyzed as that of a one-half fractional replication of the possible 2^4 experiment (Cochran and Cox, 1957). Since the sequence interactions may be assumed to be small, and the primary effects of interest (F, M, and FM) were confounded with sequence interactions, little difficulty of interpretation would be expected to arise from the confounding.

As originally proposed, (see Shrage, et.al., 1969), this experiment was intended to be a half-plaid Latin square involving the additional comparison of overt and covert responding. However, in implementation of the design the instructions relating to form of responding were accidentally omitted, and media inadvertently reassigned to groups so as to result in the design as presented above.

In the analysis of variance, the score for one student in Groups 1 and 4 in CPT 7 were replaced by least squares estimates, requiring the deduction of two degrees of freedom from the Student X Unit in Half X Groups source of variance. In one case the student received instruction under the wrong experimental condition in Unit 7, and in the other case the student took the progress checks prior to the CPT, and failed to receive instruction on one segment of Unit 7.

Materials and Procedure

The segments involved in this experiment were 2.2 - 2.5, 3.1 - 3.4, 5.7 - 5.10, and 7.1 - 7.4, and the media used were audiotapes and videotapes with panelbooks (see Table 2).

The lecturer for the videotapes was an active Naval officer with teaching background. With the aid of a teleprompter, the lecturer presented all material verbatum from a prepared script. All videotapes were prepared at a commercial television station. Commercial quality quad video tape recorders which provided broad editing capability unavailable in one inch VTR format were used. Two cameras were used in taping the lecture to allow for integration of a series of visuals (charts, photographs, drawings, etc.). Additionally, key points of content were superimposed on the screen during the program.

Two basic modules were prepared; they varied in the response demand frequency (RDF) dimension of the presentation design.

One module of the videotape lecture was characterized by high response demand frequency. At appropriate points in the lecture, the lecturer referred the students to a numbered question in the panelbook. Students read the question and recorded their answers on an answer sheet. Sufficient time was allowed for students to respond so they did not have to turn off the VTR. The number of questions asked ranged from 15 to 22. The low response demand frequency module of the videotape lecture was developed by simply editing out all but three of the lecturer's references to questions in the panelbook. All other elements of the videotape lecture remained the same.

The lecturer for the audiotapes was a commercial radio announcer. The lecturer presented all materials verbatim from a prepared script. In the segments involved in this experiment, the scripts for audio and video presentations were identical. The high and low response demand frequency modules were developed exactly as they were for the videotapes.

All audiotapes were developed in a commercial recording facility. In developing the audio presentations, standard recording tapes were used. For student use, tapes were transferred to C-60 and C-90 cassette cartridges.

All charts, photographs, drawings, etc., accompanying the audiotape lecture were presented in a panelbook. (It should be noted that for videotape modules, these graphics, etc., were presented on the VTR.)

RESULTS

Mean student performance as a function of media, response demand frequency, and type of test, is presented in Table 5. Performance under audiotape conditions may be seen to be consistently superior, although the magnitude of the difference is small. There is also a slight tendency for performance to be better under high response demand frequency conditions, especially in the Type II means. The analysis of variance, (see Appendix B) however, showed that only the effect of media was statistically reliable in the Type I data, $F(1, 40) = 4.24$, $p < .05$. No other effects of primary interest were found to be statistically significant.

Among the secondary effects, large differences between halves and units in half (see Appendix C) were found to be highly significant in every case but one. Significant interactions of Sequence X Frequency X Unit in Half were found for Total CPT scores and Type I scores. On inspection, these latter effects were found to be insufficiently large to disturb the primary pattern of results pointed out above.

Table 5

Unweighted Mean Percent Correct Responses

for Conditions of Instruction in Experiment I

	Medium			
	Audiotape		Videotape	
	Response Demand Frequency			
Test Score	High	Low	High	Low
Total CPT	65.60	64.00	62.99	61.08
Type I	65.42	64.54	62.12	62.02
Type II	65.41	62.78	63.89	57.44
Progress Check	75.02	74.61	73.02	72.22

DISCUSSION

The major finding of importance in the present experiment was that RDF had no substantial effect on performance with transient media. Unlike previous experiments with persistant media, reviewd by Anderson (1967) and Gagne and Rohwer (1969), the simple insertion of questions was not sufficient to improve performance. In view of other studies (Gropper and Lumsdaine, 1961; Abbey, et. al., 1963) which showed effects of combined questioning and feedback in transient presentation conditions, the present results suggest that feedback is essential to the improvement in performance. It also may be suggested that the importance of feedback in transient conditions may be a consequence simply of the additional presentation in the form of question and feedback, of important information which may have been missed at its first presentation.

While the present experiment showed a small difference in performance favoring the audiotape medium, this finding is not of great importance by itself. The effect may not represent intrinsic properties of these media since the lectures differed in the two presentations. However, this finding does suggest that differences in performance can be produced by media variables even when identical presentations are delivered by the media. Thus, the present results do not support Tosti and Ball's hypothesis. The two lecturers, of course, are to be considered a media variable rather than a presentation design variable.

The present media results do indicate that audiotape presentation should be chosen over videotape, particularly in view of the substantially greater cost of preparing videotape, except in those cases where visual demonstration requiring movement is essential for instruction. In other cases requiring visual presentation, the persistent panelbook form of presentation should prove adequate and provide substantial cost reduction.

VI. EXPERIMENT II

This experiment was concerned with the effects of response demand frequency (RDF) and form of response demand (RD) on learning in the persistent medium of linear programmed text. The hypothesis tested was that high RDF would be more effective than low RDF whether overt-selected, overt-spoken, or covert RD was involed.

Literature relating to effects of RDF was discussed in relation to experiment one. It was pointed out that a number of studies using the persistent narrative text medium have demonstrated both general and specific attentional effects of overt responding. However, experiments using linear programmed text are not so clear as to the effects of conditions of responding. As summarized by Anderson (1967), the large majority of studies show that students required to make overt constructed responses do not learn anymore than students who "think" the answers, or read the program with the responses filled in. It should be noted that the comparison of constructed response and "thinking" represents a comparison of overt and covert RD conditions according to the Tosti-Ball (1969) model.

It is not certain whether programs with responses filled in should be considered high or low RDF conditions. It may be that a rhetorical question with the answer provided nevertheless functionally equivalent to the same question with no answer. This might especially be expected to be the case with well-programmed linear text with low error rates. Low RDF conditions are perhaps better represented by programs in which the questions are rephrased into direct statements. This form of low RDF condition is investigated in the present experiment.

Although most studies of form of RD have found no differences, Anderson (1967) points out that studies by Holland (1965), Holland and Kent (1965), and Krumboltz (1964) suggest that overt constructed responses are facilitating if they are relevant to the most critical material to be learned. Experiments by Williams, (1963, 1965) also show that overt-constructed responses aid in the learning of the response itself.

Anderson (1967) hypothesizes that overt-constructed responses should be required when the student is expected to learn novel or technical terms, but if recognition is required or the student is already capable of emitting the response, then a multiple choice response, thinking the response, or just reading may do as well. A few studies with overt-selected (multiple-choice) responding (Williams, 1963, 1966) show that such a condition is not inferior to overt-constructed response.

The present experiment compares overt-selected and covert RD conditions, together with an overt-spoken condition which had not been investigated previously. The main purpose of the RD variable in this study is to provide an assessment of RDF-RD interactions, and a basis for the study of the relationship of RD conditions to the student variables in subsequent regression analysis. While no effects of RD conditions are expected by themselves, the effects of RDF may possibly be found to depend on the RD condition. Furthermore, certain types of students may do better using one RD form rather than another.

METHOD

Design and Analysis

The basic plan of the experimental design is shown in Table 6. The students were randomly divided into six groups, each group receiving a particular sequence of 3 RD conditions in CPT 4, 6, and 10. Groups 1, 2, and 3 are trained using high RDF materials in all CPT units while groups 4, 5, and 6 use low RDF materials.

It should be pointed out that the sequences for the first three groups are arranged to form a 3 x 3 Latin square, and the sequences for the second three groups involved the same arrangement. Since the RDF conditions are assigned to independent groups, this variable is evaluated on a between-group basis. The use of two Latin squares, each with a particular level of RDF, permits the examination of the interaction with RDF of each within-student variable forming the Latin square. In particular, the RD and RDF x RD effects of interest in the present experiment are both within-student effects. Since effects of RDF are predicted, the evaluation of this effect with lesser precision as a between student variable was regarded as conservative.

Table 6

Plan of Experiment II

Group	Response Demand Frequency	Cumulative Posttest Unit		
		4	6	10
1	High	Overt-selected	Overt-spoken	Covert
2	High	Overt-spoken	Covert	Overt-selected
3	High	Covert	Overt-selected	Overt-spoken
4	Low	Overt-selected	Overt-spoken	Covert
5	Low	Overt-spoken	Covert	Overt-selected
6	Low	Covert	Overt-selected	Overt-spoken

It may also be noted that the present design permits the examination of interactions between units and the variables of primary interest. Failure of the unit interaction to appear as significant would tend to support the generality of the effects of the primary variables across unit and content. In addition, such results would indicate that effects of the sequence of RD conditions were negligible.

The score of one student in CPT 6 was replaced by a least-squares estimate since he had proceeded through the wrong module.

Materials and Procedure

The segments involved in this experiment were 4.1 - 4.3, 5.4 - 5.6, and 8.1 - 8.3, and the medium used was linear text (see Table 2).

Linear programed texts were developed by the RULEG and EGRUL methods of programing (Rule-example; example-rule). These are essentially programing methods of presenting a rule (definition, principle) and having the student identify an example of the rule (from 2, 3, or 4 choices), or presenting an example and having the student identify the rule or principle which is depicted in the example (Markle, 1964). Variations of the RULEG-EGRUL method which were used are EG-EG and RUL-RUL.

It is important to note that although confirmation of responses is ordinarily an important part of programed instruction, confirmations were deleted in the first implementation in order to obtain valid data on student frame responses. The first three modules of linear text were presented in standard format; i.e., there was a question for every information frame. This defined the high response demand frequency (HRDF) dimension. The only difference in presentation among these modules was in the form of response required of the student.

In module 1 (HRDF-written RD), students were instructed to respond to each frame by writing their selection (A, B, C, or D) on the frame answer sheets. In module 2 (HRDF-spoken RD), students were instructed to respond to each frame by speaking their selection (A, B, C, or D) into a tape recorder. In module 3 (HRDF-covert RD), students were instructed to read each frame question and think the answers to themselves.

Modules 4, 5, and 6 of the linear text covered the same material as the first three modules, but the frequency of response demand varied within the presentations. Instead of asking a question for every frame, questions were asked for every second or third frame. In "no-question" frames, examples or principles, which would be deleted when questions were deleted, were reworded in statement form; e.g., instead of asking, "Which of these situations best exemplified principle X?", the frame was followed by a statement, such as "An example of principle X is. . ." Modules 4, 5, and 6 differed from each other in the form of response demand similar to modules 1, 2, and 3.

RESULTS

As expected, high RDF conditions were found to be consistently superior to low RDF conditions. However, the means presented in Table 7 show the differences produced by this variable were not large, ranging from about 6% of the Total CPT scores to less than 3% in the Progress Check. The analysis of variance (see Appendix D) only showed a significant RDF effect in the analysis of Total CPT scores, $F\ (1,\ 38) = 4.25,\ p < .05$.

There were no interactions between RD and RDF found to be statistically significant. However, a significant main effect of RD was found in the analysis of progress checks, $F\ (2,\ 75) = 3.86$, $p < .05$. As the means presented in Table 8 show, there was little difference among RD conditions for Total CPT, Type I, or Type II scores. Performance was slightly better with overt-selected responses.

Among the secondary findings, the differences among units was found significant in every analysis, but there was no evidence of significant interaction between units and the primary variables.

TABLE 7

Unweighted Mean Percent Correct Responses for

Response Demand Frequency Conditions in Experiment II

	Response Demand Frequency	
Test Score	High	Low
Total CPT	66.78	60.63
Type I	70.28	66.03
Type II	62.84	58.19
Progress Check	77.33	74.61

TABLE 8

Unweighted Mean Percent Correct Responses for Response Demand Form Conditions of Experiment II

	Response Demand Form		
Test Score	Overt-written	Overt-spoken	Covert
Total CPT	63.65	65.06	64.87
Type I	66.69	69.14	68.85
Type II	60.30	60.53	60.72
Progress Check	78.50	74.11	75.31

DISCUSSION

The results of the present experiment are not inconsistent with previous literature on the effects of RD and RDF. Like many previous experiments high RDF had no significant facilitating effects when measured by direct test of the instructional objectives such as those represented by the progress check items. The present findings indicate that high RDF may have its primary effects on generalization and problem solving items which have general content validity but no specific criterion reference, as is the case with the Type I and Type II items making up the Total CPT score. This may represent the kind of general facilitary effect of questioning previously reported for narrative text.

It should be noted that there is no contradiction in finding significant results for Total CPT scores while not finding significant differences with respect to the components Type I and Type II subtests. Since the Total CPT test is a longer test than either of its components, the consequent increase in reliability of the Total CPT scores produces a reduction in variability which increases the likelihood of finding significant and true difference between conditions.

The finding of a significant superiority of the overt-selected RD conditions on progress checks does not violate the general findings that form of RD make little difference in achievement. The obtained difference was quite small (only 3% better than covert) and probably represent the same kind of specific training effect as that reported by Williams (1966). Williams found that while there was no overall difference between overt-constructed and overt-selected responding, performance was better when the response form required on the test was the same as that used in training regardless of which form was used. Since the test required overt-selected responses it is not surprising to find that training with overt-selected responses is somewhat superior on direct criterion-referenced items, many of which may have appeared in the same or similar form of training. In general, the present findings support the conclusion that high RDF presentations using the same RD forms required on the test are usually to be prefered in the preparation of linear programmed text materials. While other conditions may on occasion be found to be substantially equivalent in performance there is no evidence with either the present findings or previous experiments that such other conditions may be found superior to high RDF with tested RD forms.

VII. EXPERIMENT III

This experiment investigated the effects of response demand frequency (RDF) and management frequency (MF) in two different media. Specifically, the hypothesis tested was that HRDF presentations with either high, medium, or low MF would be superior to a low RDF, low MF presentation. The Computer Assisted Instruction (CAI) and Audiotape-Intrinsically Programmed Booklet (AT-IP) media selected for this experiment provided capability for the manipulation of MF, and enabled a test of the generality of the findings concerning RDF and MF.

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 efficient linear programs have:

(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 reviewing this literature, Anderson (1967) has suggested that overt constructed responses are useful primarily in learning unfamiliar technical terms. He also concludes that knowledge of results serves as corrective feedback rather than a reinforcer, and is largely unnecessary when the student responds correctly. The implication is that knowledge of results may be omitted from well-programmed material with low error rates; i.e., that learning can be assured by means built into the stimulus materials themselves. As discussed in relation to Experiment I, the effect of responding may be essentially in directing and maintaining attention. 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 (1969) 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, 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 performance by the students.

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 possibility 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.

Several studies reviewed by Anderson (1967) have provided feedback on an intermittant schedule, usually with facilitating effects. However, students in the control groups receiving feedback have usually had the opportunity to copy the responses without paying attention to the instructional material, a factor which undoubtedly had deleterious effects. Thus, the independent effects of RDF and MF have not been adequately separated. In the present study, this separation is made, based on 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.

METHOD

Design and Analysis

The basic plan of the experiment is presented in Table 9. The students were randomly divided into four groups, each group receiving one of four combinations of RDF and MF conditions of instruction in both CPT Units 5 and 13. The AT-IP media was employed for the presentations in Unit 5, and CAI for presentations in Unit 13. This arrangement represents a "mixed" design, with RDF-MF conditions as the between student variable, and unit-media combinations as the repeated measure variable.

Since differences in performance between CPT Units 5 and 13 could be expected from differences in difficulty of content or test items, or quality of programming, in addition to differences between media, the present design does not provide a meaningful direct comparison between media. The design does provide an estimate of the effects of combinations of RDF and MF averaged over both media. In addition, the interaction between RDF-MF conditions and unit-media combinations could be examined to determine if the effects of the presentation variables were comparable in both media.

The analysis of variance followed standard procedures for "mixed" design as modified for an unweighted means analysis (see Winer, p. 374). As a result of computer

TABLE 9

Plan of Experiment III

Group	Cumulative Posttest Units	
	3 (AT-IP)	5 (CAI)
1	High RDF - High MF	High RDF - High MF
2	High RDF - Medium MF	High RDF - Medium MF
3	High RDF - Low MF	High RDF - Low MF
4	Low RDF - Low MF	Low RDF - Low MF

breakdown in the CAI presentation for 10 students, and the loss of CPT data for 2 students in AT-IP conditions, and 1 student in CAI, the analysis of progress check data was based on 34 students and the analysis of CPT data on 31 students. Groups 1, 2, 3, and 4 had 5, 7, 11, and 8 students, respectively, in the case of progress checks, and 6, 9, 11, and 8 students, respectively, in the case of CPT data.

Materials and Procedure

The segments involved in this experiment were 4.4 - 4.7 and 12.1 - 12.4. The medium used in segments 4.4 - 4.7 was audiotape with an intrinsically programed booklet. Computer-Assisted Instruction (CAI) was used in segments 12.1 - 12.4.

As originated by Norman A. Crowder, the intrinsic programing technique consisted of routing a student through a "scrambled" text on the basis of his response. Each response directed him to a different page of the text; thus, the student could not read through directly and sequentially.

Combining the intrinsic programing technique with an audiotape was a WLC innovation. In this teaching mode, the information was presented via the tape. While the student listened to it, he also looked at a summary page in the text which contained the precis of what he was hearing. He then stopped the tape and followed the instructions at the bottom of the summary page, directing him to a page containing a question which tested the information given on the tape and summary. Each response to the test item referred the student to another page which informed him of the accuracy of that response. Thus, the student would select the alternative which he thought was correct, turn to the page indicated for that alternative, and find out if he had made a correct selection. If he had selected the correct response, he was instructed to go on to another summary page which he read while listening to the next audio portion. If his response was partially correct or incorrect, he was either told the nature of his error and instructed to proceed as described above, or he was instructed to return to the summary or question page to study the information again and select another alternative. This process of interaction between tape and text continued throughout the segment.

The tape, which contained the content of the segment, remained the same throughout the four modules. The text differed as follows:

In module 1 a question was asked for each informational frame (HRDF). Based on his response, the student was always branched to a page where his answer was discussed and confirmed or rejected (high management frequency).

In module 2 a question was asked for each frame (HRDF), but the student was branched on the basis of his response for only 50% of his responses (medium management frequency).

In module 3 a question was asked for each frame (HRDF), but the student was never given feedback nor confirmed; i.e., he was never branched as a result of his selection. He simply went on to the next question (low management frequency).

In module 4 only three or four questions were asked throughout the programed sequence (LRDF), and the student was not branched on the basis of his responses to the three or four questions (low management frequency).

All of the three components of the CAI 1500 system (CRT, audio, and image projector) were utilized in the implementation of the four modules. These modules exactly paralleled the four modules used in the audiotape-IPB segments. The following pattern was followed in developing the CAI materials. The informational frames were presented on the CRT screen and image projector. The questions, which were often situations in which the student had to decide the best course of action, were presented: 1) on the audio, where the situation was described; 2) on the image projector, where pictures of the situation were presented along with the audio; 3) on the CRT screen, where the student was asked to select an answer from 3 to 5 choices. The student's selection, accompanied by feedback, was displayed on either the CRT screen or the audiotape, and occasionally on the image projector. This feedback consisted of the reason(s) why the selected answer was correct or incorrect.

RESULTS

Mean performance in Experiment III as a function of RDF-MF conditions and type of test is presented in Table 10. Little variation in performance is to be seen for Type I scores or progress checks, while performance in the high RDF-high MF condition appears to be substantially inferior to the remaining conditions for Type II scores, and consequently, for Total CPT scores as well. The results of the analysis of variance confirmed the reliability of this latter effect in the case of the Type II scores; $F(3, 27) = 3.40$, $p < .05$ (Appendix F).

Significant differences among unit-media combinations were also found in the Type II and Progress Check analyses, but no interactions of this variable with RDF-MF methods were found to be statistically significant (Appendix G).

Table 10

Mean Percent Correct Responses for

Conditions of Instruction in Experiment III

	Response Demand Frequency			
	High	High	High	Low
	Management Frequency			
Test Score	High	Medium	Low	Low
Total CPT	63.25	68.57	66.14	68.59
Type I	65.76	66.13	64.43	67.77
Type II	60.38	70.97	67.69	69.34
Progress Check	86.74	84.04	83.00	82.68

DISCUSSION

The finding of inferior performance on Type II items under high RDF-high MF conditions is rather surprising, particularly in relation to the CAI medium. The primary claim for the effectiveness of intrinsic (branched) programming (Crowder, 1959), and especially CAI, is based on the additional capability for individualized management of instructional sequences based upon responses to strategic questions. The present results suggest that the high-frequency utilization of management may tend to interfere with the acquisition of skills required for problem-solving on the basis of learned information. Furthermore, no substantial advantage is to be obtained from even a moderate use of management, since the remaining conditions were found to be essentially equivalent in effectiveness.

The basis of inferior performance with HMF may possibly be found in the fact that branched programming and CAI management involves the presentation of additional comments, examples, and questions which may go somewhat afield from the main line of instruction as well as delay progress through the primary sequence of instruction. This may have the effect of interfering with the integration and interrelating of diverse knowledge and principles required for successful problem solving performance on Type II items.

The equality of performance under the remaining conditions is still puzzleing, even given the interpretation above. It should be pointed out, however, that this result may depend on the fact that no confirmation was given in the present experiment to minimize uncontrolled self-management by the student. It may be that the optimal properties of moderate levels of MF in these media would be revealed if confirmation were provided.

VIII. EXPERIMENT IV

One major goal of the preceding three hypotheses was to demonstrate that students will generally perform better on presentations with high RDF. Numerous exceptions to this principle, however, have been a source of many criticisms recently leveled at linear programmed 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, 1969), 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 programmed instruction.

One of the chief complaints 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 RDF presentations, then criticism of PI for failure 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.

Experiment IV is a direct test of the effectiveness of the syndactic-text procedure. It is hypothesized that student's unable to achieve satisfactory levels of performance after practice on a low RDF presentation, will show superior levels of performance if remediated on a high RDF presentation, as opposed to an additional low RDF presentation or no remediation. If results support this hypothesis, one means of adapting the Tosti-Ball (1969) model to individual differences in instructional needs is suggested which may have considerable generality. For example, a low RDF audiotape presentation may be found ineffective for some portion of the students, who might then be remediated on high RDF linear text. The entire instructional sequence might prove more efficient than either audiotape or linear text would be if used alone.

METHOD

Design and Analysis

The plan of this experiment is shown in Table 11. The students were randomly divided into six groups, each group receiving a different sequence of three experimental methods in CPT Units 3, 8, and 12. In all conditions, the students received the same syndactic-text summaries; the indicated methods of instruction refer to the type of supplementary material used when an inadequate level of performance was achieved on the quiz following a summary.

It may be noted that all six possible sequences of three conditions are used in this design. The sequences are arranged to form two different 3 X 3 Latin Squares, one square involving Groups 1, 2, and 3, and the second square involving Groups 4, 5, and 6. The difference between these squares may be seen in the reversed sequence of conditions in Units 8 and 12. The consequence of this reversal of sequence is that different components of the Method X Unit interaction are confounded with group differences in each square. Since the groups are randomized, significant differences among groups may legitimately be interpreted as arising from the confounded Method X Unit interactions. However,

TABLE 11

Plan of Experiment IV

Group	Cumulative Posttest Unit		
	2	8	12
1	High RDF	Low RDF	None
2	Low RDF	None	High RDF
3	None	High RDF	Low RDF
4	High RDF	None	Low RDF
5	Low RDF	High RDF	None
6	None	Low RDF	High RDF

the remaining components of this interaction are independently estimable from the residual variation in each square, after removal of Unit, Method, Group variation. Thus a second, and more precise, test of the Method X Unit interaction is obtained on a within-student basis.

The present design also permits the examination of Square X Unit and Square X Method interactions, which in this case must be interpreted as arising from effects of the sequence of methods.

The analyses were conducted according to the procedure described by Winer (1962, p. 549) as modified for an unweighted means analysis. The score of one student in CPT 12 was replaced by a least-square estimate, since the student had remediated the segments in this unit prior to taking the CPT test.

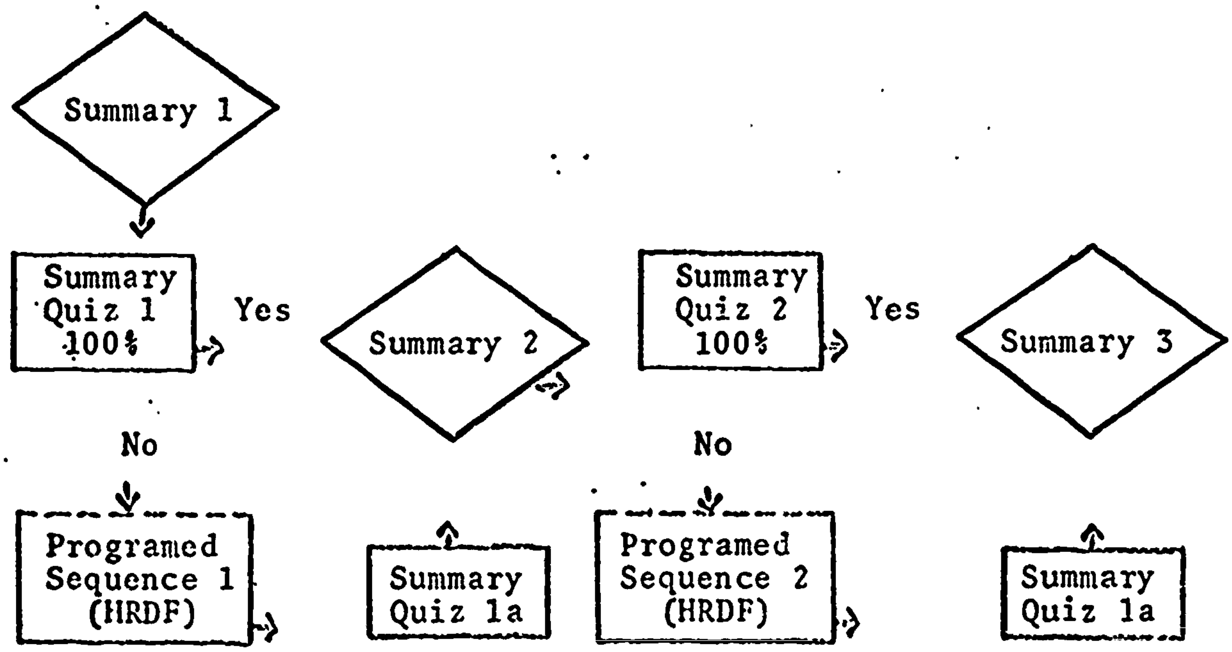
Materials and Procedures

The segments involved in this experiment were 2.6 - 2.8, 6.1 - 6.3, and 11.1 - 11.3, and the medium used was syndactic text (see Table 2).

A syndactic text is essentially a series of linear programed frames each preceded by a brief but complete summary of the information presented in the frames. Students worked through the syndactic text by reading the first summary statement and taking a summary quiz of 5 to 8 questions. If the student answered all summary quiz questions correctly, he read the second summary, took summary quiz 2, etc.

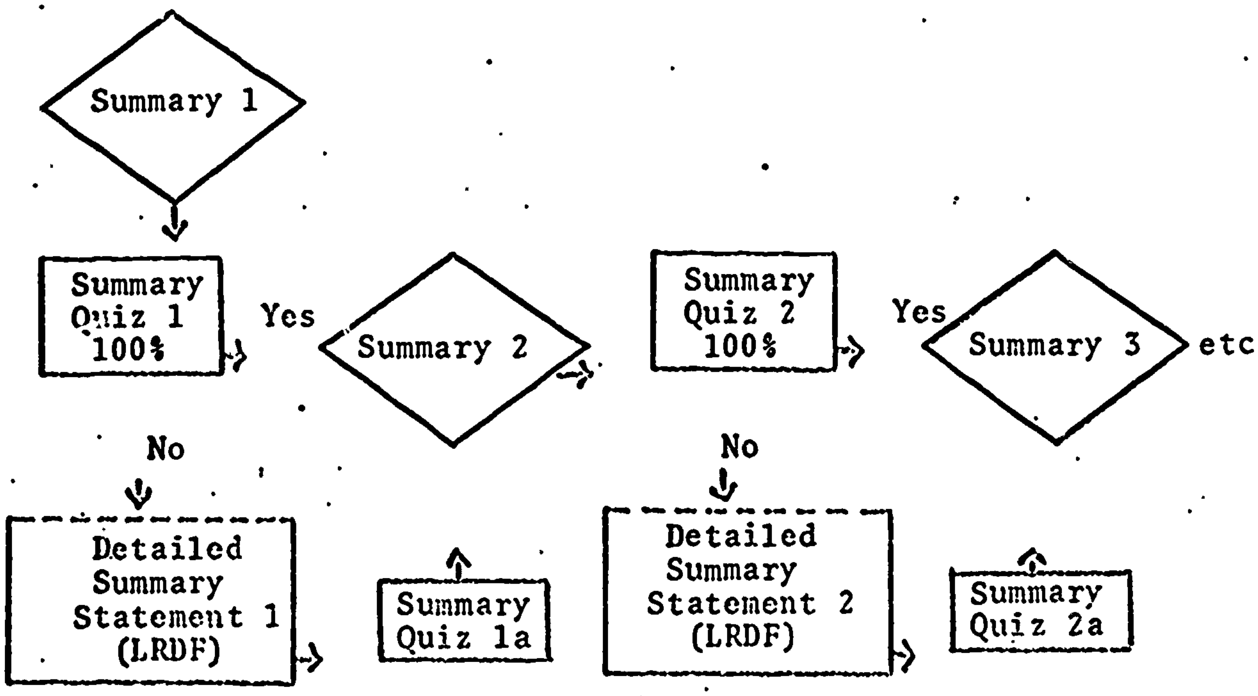
The student who incorrectly answered one or more questions of a summary quiz was required to study the linear programed sequence was identical to the linear text discussed on page 21. It was developed by the RULEG or EGRUL method of presenting small bits of information accompanied by examples of the concepts being taught. At the end of the programed sequence, the student retook the summary quiz. Regardless of his performance when he retook the summary quiz, he went on to the next summary statement and repeated the procedure. Non-research segments of syndactic text were implemented according to the procedure given above.

The first module of research segments of syndactic texts was presented in the same manner as non-research segments. It was characterized by high response demand frequency (HRDF). Modul? 1 is depicted in chart from as:

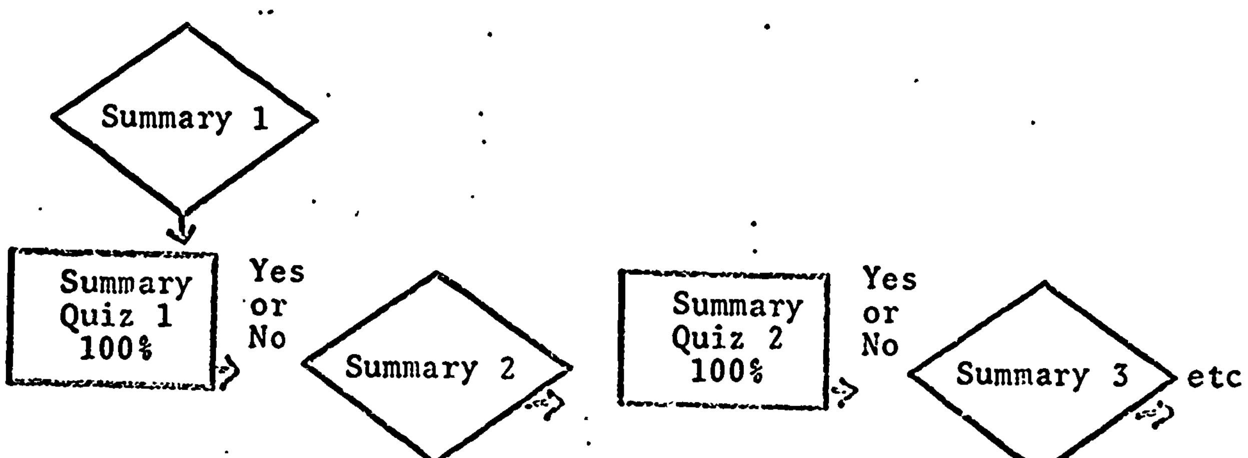


The second module of syndactic text was identical to module 1 except that it was characterized by a low response demand (LRDF) program. Instead of having RULEG question frames, examples were simply given in statement form. This sequence of statements was referred to as a "detailed summary statement." There were no questions asked in the detailed summary statement.

Module 2 is depicted in chart form as:



The third module was represented by summary statements alone (no remediation, no RDF). The student read a summary statement, took the summary quiz, and proceeded to the next summary statement regardless of his score on the summary quiz. Module 3 is depicted in chart form as:



RESULTS

The present experiment failed its intended purpose since it was found that students passed the summary quizzes in the large majority of cases. Frame sequence data show that 73% of the summary quizzes were passed, permitting the student to proceed directly to the next summary. Thus, the students did not contact the experimental variation in remediation method contained in the supplementary material with any frequency. As a consequence, performance was closely similar under all conditions, as the means presented in Table 12 show.

Only differences among units and one Sequence by Unit interaction were found statistically significant in the analyses of variance (see Appendix H).

The means for unit by method combinations are presented in Appendix I.

DISCUSSION

The primary conclusion which may be drawn from the present results is that well-prepared syndactic-text summaries are surprisingly effective for students at the level of the Naval Academy midshipman. While no inference may be made concerning the best type of presentation to include as a back-up to the summaries, such results have the important practical implication that many highly designed and controlled instructional procedures may be relatively inefficient both in relation to developmental costs and student learning time. It is suggested that the syndactic text format deserves increased utilization and critical analysis as an educational method for the presentation of content materials similar to those involved in the Leadership Management course.

IX. EXPERIMENT V

In each of the four previous experiments 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 embodded 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 achievement levels may be significantly improved if the student is allowed to participate in the process of instructional management. As Tosti (1969) 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.

The old saw that "the best way to learn is to teach" would suggest that student involvement in instructional management of other students can be of considerable value to the student manager's own educational attainment. Such effects form part of the rationale for the time-honored seminar system. Although, the belief that such effects may exist have had wide currency for a considerable number of years, it is only recently that empirical evidence has begun to be developed to support this notion. Cloward (1967), for example, reports data showing marked increases in the reading ability of students who served as tutors in reading with other students.

Many people interpret the notion of an effect of teaching on learning as a direct one, i.e., that teaching particular content material increases the learning of that same content. The most impressive study of the effect of teaching on learning, however, shows that the influence may be much broader and fundamental than supposed, representing a transfer from teaching to subsequent learning of new content. Long (1970) had "teaching-students" present paried-associate lists and response feedback to other students. Subsequently learning of new lists by the "teaching-students" was markedly facilitated relative to controls. Interestingly, "teaching-students" who taught non-responding students were not found to show this facilitation, implicating the feedback to the teacher of student behavior as a critical factor in the facilitating effect.

The present experiment is concerned with evaluating a procedure designed to produce the "student-teaching" effect in the Naval Academy context. This procedure is built upon the common occurrence of joint-studying among peers, but provides that students will systematically adopt student and teacher roles. The hypothesis is that student's in the experimental "peer-monitor" system involving student and teacher roles will show faciliation of learning relative to controls who continue with the usual individualized instructional system.

METHOD

Design and Analysis

The students were randomly assigned to the two groups. Two students originally assigned to the peer-monitor group were dropped due to their inability to schedule time together. One of these students was the fastest in the class, the other was the slowest. A one way analysis of variance was employed in analyzing the data.

Materials and Procedure

The segments involved in this experiment were 8.4 - 8.6, and the medium used was the learning activity summary (LAS).

A learning activity summary is similar to a traditional textbook or bibliography approach to learning. It is a technique very often used in college and graduate seminars to put the responsibility for structuring learning and achieving objectives on the student. A learning activity summary was composed of three parts: 1) an overview of the segment, 2) behavioral objectives for the segment, 3) a bibliography of source material that was related to each of the objectives. Students worked through an LAS segment by reading the overview and objectives and studying text materials which were related to each objective. Text materials were either select pages in published text books or supplemental handouts.

The student studied all text material until he felt he had mastered the objectives and could pass the progress check. If he did not achieve 80% of the objectives on the progress check, he remediated by re-studying the prescribed text material. In module 1 in this medium, students initially studied alone and then worked in pairs to prepare for the progress check. In module 2 students studied individually.

RESULTS

The primary findings of the experiment are presented in Table 13. There is a tendency for better performance in the Peer-Monitor Group on Type I items, but the opposite finding may be seen for Progress Check performance. Statistical analyses showed none of the differences in mean levels of performance reached accepted levels of significance (Appendix J). Examination of the intra-group variability, however, showed that the variance was significantly smaller in the Peer-Monitor Group on the Type II data, $F\ (23, 17) = 2.53,\ p < .05$. Apparently, student interchange in the Peer-Monitor pairs led to a greater homogeneity of opinion on problem-solving questions without influencing proficiency.

Perhaps the most significant finding of this experiment was that the Peer-Monitor procedure as presently constituted was almost universally regarded with distaste by the students. Apparently difficulties in scheduling, and feelings that the interaction was superfluous following individual studying detracted from the potential effectiveness of the procedure.

TABLE 13

Means and Standard Deviations of

Percent Correct Responses in Experiment V

	Peer Monitor Group		No Peer Monitor Group	
Test Score	M	s^2	M	s^2
Total CPT	68.5	9.7	64.3	11.1
Type I	77.3	6.0	70.4	6.7
Type II	58.4	4.9	57.2	7.8
Progress Check	75.5	5.4	79.1	4.8

DISCUSSION

The present experiment demonstrates that student interaction in the particular form attempted here, does not have an inevitable facilitating effect on performance. If the "teacher-student" effect is to be demonstrated some more palatable and logistically-feasible procedure remains to be designed.

X. MEDIA COMPARISONS

As has been repeatedly mentioned, Experiments I through IV were primarily designed as studies of presentation design factors within particular media. Only Experiment I incorporated a direct comparison between media, the differences in media in Experiment III being confounded with differences between units. However, a general comparison among media in different experiments is possible if certain specific assumptions concerning the CPT units can be satisfied.

First, it must be assumed that differences in performance among experimental modules are either negligible, or if present, are typical of the extent of performance variations which might be incountered among CPT units developed independently for the same content in that medium. If this assumption is satisfied, then mean performance for a particular unit averaged over different modules are not systematically biased either upward or downward by the effects of the modules used in the unit. The results reported previously tend to support this assumption, since the differences among experimental modules were found to be small in magnitude even when statistically significant.

Second, it must be assumed that CPT units in a particular media involved random (or at least representative) samples from population of content, skill of preparation of instructional materials, and difficulties of test items. If this assumption is correct, the mean performances on units within media may be regarded as a representative sample from a population of potential unit means which are typical of implementations of the media at the present stage of development and revision of materials. It may be noted that such an assumption concerning the sampling of content is patently false in view of the systematic selection of media which was employed. However, the media were assigned in a roughly balanced fashion throughout the course, which suggests that the units prepared in different media are reasonably representative of the types of content involved in the course. The writers of materials and test items were similarly varied in assignment.

The third assumption required is that the population distribution of means for units within media are normally distributed with equal variances among media population. The assumption of normality is assured to be approximately correct in the present case by appeal to the "central limit theorem" (see Hays, 1965) since each mean is based on a substantial number of scores. The assumption of homogeneity of variance also seems viable on inspection of the data.

Given the assumptions above, the means for units within media may be treated as a sampling element forming the basis of a simple one-way analysis of variance comparing media. Since significant differences among units were demonstrated in all previous analyses of experimental data, it is obvious that the present comparisons among media are made at a considerably reduced level of precision. Nevertheless, any gross differences in media effectiveness beyond the level of normal variation among units in media would be detectable in the present analyses.

RESULTS

Performance averaged for the CPT units of six media are presented in Table 14. It should be noted that the videotape medium is omitted from this comparison since it was demonstrated to be somewhat inferior in the analyses of Experiment I. The means presented are based on from 1 to 4 units, involving from 3 to 16 segments. It should also be pointed out that performance on some segments which were not assigned to CPT units and which were not used for experimental purposes are not included in these averages.

The major impression to be gained from inspection of Table 14 is that there is no entirely consistent pattern of superiority or inferiority for any medium across all types of test scores, nor are there any massive differences in performance, although differences among the extreme cases appear substantial. Caution should be exercised in interpreting differences between the extremes, since it is a well-known statistical phenomenon that substantial differences inevitably will occur entirely by chance among the extreme values.

The results of the analyses of variance (see Appendix K) were highly consistent in showing no significant differences among media for any test. The means for media by cumulative Posttest Unit for each type of test are given in Appendix L.

TABLE 14

Unweighted Mean Correct Responses for

Media, Averaged over Modules Segments and Units

Number of Units	Number of Segments	Test Score			
		Total CPT	Type I	Type II	Progress Check
		Audiotape Lecture with Panel Booklet (AT-PB)			
4	16	63.42	63.52	62.38	74.63
		Linear Text (LT)			
3	9	64.52	68.16	60.52	75.97
		Syndactic Text (ST)			
3	9	60.20	63.46	56.05	81.12
		Audiotape and Intrinsically Programed Booklet (AT-IPB)			
1	4	64.96	68.00	61.62	81.39
		Computer Assisted Instruction (CAI)			
1	4	68.31	64.28	72.59	86.84
		Learning Activities Summary (LAS)			
1	3	66.37	73.88	57.79	77.32
		Unweighted Media Average			
13	45	64.63	66.88	61.82	79.54

DISCUSSION

The present findings of no statistically reliable differences among media indicate that the substantial differences which were observed can reasonably be attributed to sampling variation arising from difficulty of content, skill in preparation materials for that content, or difficulty of test items. Such results do not provide strong support for the Tosti-Ball hypothesis of no intrinsic media differences, because the present media comparisons were made at a relatively gross level of experimental precision. Comparison among media with the same kind of control and precision devoted to comparison of presentation variables in the present studies may well have demonstrated numerous reliable differences among media, as was the finding in Experiment I. However, these results do support the conclusion that differences among media, if they exist, are no more substantial nor important than variation among units in the same media, resulting from variation in content programming skill, or test items. Thus it would seem that the problem of media selection should deserve no more attention from the educational technologist than he is willing to devote to these other sources of variation in student achievement.

XI. CONCLUSIONS AND IMPLICATIONS

This research effort was directed at answering the central question of whether variations of conditions of instruction in the presentation design domain as posited by Tosti and Ball (1969) are of greater importance than variations in the media domain. Tosti and Ball's position has not received strong experimental support from this series of studies. Indeed, in Experiment I where the presentation design was held constant, a significant difference was found between media. Although this does not necessarily refute the basic assumptions of the Tosti and Ball model, it does indicate that other variables relating to the production of instruction particularly via transient media (videotapes and audiotapes) should be considered.

In investigating the question of whether substantial effects of student achievement could be produced by manipulation of presentation variables over large segments of instruction, Experiment II provides the most conclusive evidence. It was found that with linear text, the high response demand frequency condition was consistently superior to the low response demand frequency condition. A facilitative effect does appear to be produced by the insertion of a significant number of questions in the instructional material. Further investigations need to be conducted with respect to this variable as implemented with transient media (see Experiment I). Although major differences were not found with respect to the form of response required of the student, the trend was that the overt-selected response condition was slightly better than either overt-spoken or covert. This finding is generally in concert with the major body of prior research that indicates that the form of response utilized within the instructional materials should be similar to that required of the student on the tests of achievement on those materials.

Perhaps the least conclusive finding relates to manipulations of management frequency in conjunction with variations in response demand frequency as indicated in Experiment III. It appears that the most efficient condition would be the use of moderate levels of management frequency in conjunction with high response demand frequency. However, an intervening variable may be the fact that knowledge of results was not provided in those cases where management was not manipulated. This may be producing variations in the experimental conditions that are not due to the major variables being investigated.

In relation to overall differences among media, the results tend to support the conclusion that differences among media, if they exist, are no more substantial nor important than variation arising from difficulty of content, skill in preparation of materials in that medium for the content, of difficulty of test items. Indeed, with respect to both media and presentation variable differences, an important factor appears to be the type of learning or task required of the student. That is, differences in conditions of instruction and media vary in relation to the acquisition of knowledge (lower level learning) as opposed to the application of that knowledge.

It should be noted here that his data should be interpreted with the knowledge that it is based on the first full scale implementation of the course materials. The course has been revised on the basis of empirical data and the research is presently being replicated with a similar population of students. A brief summary of changes to the experimental hypotheses for the replication being conducted this fall are given below.

In the first implementation of the course, confirmation or knowledge of results was not used. It was felt that more reliable data needed for revision of materials could be obtained without confirmation. For the fall semester replication, confirmation within all instructional sequences has been added.

In Experiment I the use of videotapes has been dropped, and the issue of confirmation is being experimentally investigated.

In Experiment II the overt-spoken condition of response demand is being replaced with a condition where the students underline the response they select.

Relatively high levels of performance were attained on the CAI materials in Experiment III. In order to determine whether this effect was due to the medium or the content programed for it, a parallel "paper" version is being prepared for both the CAI segments and the audiotape-intrinsically programed booklet segments.

With respect to Experiment IV, it was found that many students did not need to proceed through the remedial sequences. The syndactic text summaries were very effective. In order to obtain a more complete test of the effectiveness of each of the remediation forms, students will be forced through the sequences and regression analysis will be used to attempt to predict when the remedial sequences will be needed and which one would be most effective for each student.

The peer interaction condition of Experiment V will not be used. The implementation of any similar but more palatable form of student interaction will be guided by the general performance and attitudes of the students prior to their working on segments 8.4 to 8.6. If implementation of an interaction condition appears, at that point, to be generally deleterious to the students progress and motivation, it will not be implemented.

REFERENCES

Abbey, et. al., in Schramm, W. S. Research on programed instruction: An annotated bibliography. Washington, D. C.: U. S. Government Printing Office, 1964. 114 pp.

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

Anderson, R. C. Educational psychology. In P. R. Farnsworth (Ed.), Annual Review of Psychology. Palo Alto, California: Annual Reviews, Inc., 1967. pp. 129-164.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives: The classification of educational goals. (Handbook I: Cognitive Domain) New York: Longmans, Green and Co., 1956. 207 pp.

Briggs, Leslie J. Handbook of Procedures for the Design of Instruction. American Institutes for Research. Pittsburgh, Pennsylvania, 1970. pp. 1-206.

Briggs

Brudner, H. J. Computer-managed instruction--conclusion. Instructional Media, Vol. 1, No. 2, May, 1969.

Campeau, Peggie L. Level of anxiety and presence or absence of feedback in programed instruction. Palo Alto, California: American Institutes for Research, February, 1966. (Mimeo.)

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.

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

Cloward

Cochran, W. D., and Cox G. M. Experimental Designs. Second Edition, New York, Wiley, 1957.

Cole, James A., and Cosgrove, William P. Leadership Management Course Student Guide (TR-6.5) Westinghouse Learning Corporation, Annapolis, Maryland, 1970.

Cole, James A., and Cosgrove, William P. Leadership Management Course Instructor Guide (TR-6.6) Westinghouse Learning Corporation, Annapolis, Maryland, 1970.

Cronbach, Lee J., Chapter 2: How Can Instruction be Adapted to Individual Differences? in Robert M. Gagne, Learning and Individual Differences, Charles E. Merrill Books, Inc., Columbus, Ohio, 1967, pp. 1-265.

Crowder, N. A. Automatic tutoring by means of intrinsic programing. In E. H. Galanter (Ed.), Automatic teaching: The state of the art. New York, Wiley, 1959, pp. 109-116.

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. pp. 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.

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

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

Evans, J. L., Homme, L. E., & Glaser, R. The RULEG system for the construction of programed verbal learning sequences. Journal of Education Research, 1962, 55, 513-518.

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.

Flanagan, J. C. Functional education for the 70's. Phi Delta Kappan, 49 (Sept. 1967), 27-32.

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.

Frye, C. H. Group vs. individual pacing in programed instruction. Oregon State System of Higher Education, 1963.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Gagne, R. M. and Rohwer, W. D. Instructional Psychology. In Farnsworth P. R., (Mc Nemar, O. and McNear, Q. (Eds.) Annual Review of Psychology. Vol. 20. Palo Alto, Califronia: Annual Reviews Ins., 1969.

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. Catalong #FS-5. 234:34041, 1967.

Gropper, G. L. & Lumsdaine, A. A. An experimental evaluation of the contribution of sequencing, pretesting, and active student responses to the effectiveness of "programmed" TV instruction. (Studies in Televised Instruction, Report No. 3) Pittsburgh, Pennsylvania: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX and American Institutes for Research, April 1961. (b).

Hays, William L. Statistics for Psychologists, Holt, Rinehart, and Winston, New York, 1965.

Holland, J. G. Research on programming variables. In Teaching Machines and Programmed Learning II (Glaser, R., Ed., National Education Association, Washington, D. C., 1965)

Holland, J. G., and Kemp, F. D. A measure of programing in teaching-machine material. J. Education Psychology, 56, 264-69 (1965)

Krumboltz, J. D. The nature and importance of the required response in programed instruction. American Education Research Journal, 1, 203-9 (1964).

Long, Karl Kilby. The Transfer From Teaching To Learning. The University of New Mexico, Albuquerque, January, 1970.

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.

Mager, Robert F. Preparing Instructional Objectives. Fearon Publishers, Palo Alto, California. 1962.

Meyers, D. E. Adjunct to self-study for aircrew refresher training under operational conditions in the Air Defense Command. Dayton, Ohio: Wright-Patterson Air Force Base, Behavioral Sciences Laboratory, 1965. (AMRL-TR-65-83).

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.

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.

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

Shrage, Jules, et. al., Leadership Management Course Research and Evaluation Plans - Part II, (TR-6.3a) Westinghouse Learning Corporation, Albuquerque, New Mexico, April, 1969.

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, W. I. & Moore, J. W. (Eds.) Programmed learning: Theory and research. Princeton, New Jersey: D. Van-Nostrand Company, Inc., 1962. 240 pp.

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

Stolurow, L. M., and Davis, D. J. Teaching machines and computer-based systems. In R. Glaser (Ed.) Teaching Machines and Programmed Learning: Data and Directions. Washington, D. C. National Educational Association, 1965, pp. 162-212.

Tosti, D. T., and Ball, J. R. A Behavioral Approach to Instructional Design and Media Selection. Albuquerque, 1969. Behavior Systems Division, Westinghouse Learning Corporation.

Williams, J. P. Comparison of several response modes in a review program. Journal Educational Psychology 54, 253-60, 1963.

Williams J. P. Effectiveness of constructured-response and multiple-choice programing modes as a function of test mode. Journal Educational Psychology. 1, 211-18, 1964.

Williams, J. P. Combining response modes in programmed instruction. Journal Educational Psychology, 57, 215-19, 1966.

APPENDIX A

APPENDIX A

SPECIFICATIONS FOR RESEARCH TEST ITEMS

A. General specifications

1. Relationship of items to content: Cumulative Post-Test (CPT) items will be prepared on the basis of content outlines and the content of "common" instructional matrrials (i.e., materials presented in each of the parallel experimental modules).
   Each CPT item must have an identifiable relationship to one or more general and/or specific headings of the content outlines; references to content outline headings will be noted on each test item submission form.
   Each item will be related to the content of as many different segments making up the Cumulative Post-Test Unit as is feasible and desirable in relation to other specifications detailed below. The entire set of items should represent an equitable distribution of content among segments of the Unit in relation to the relative amount of content covered by the outlines and the materials of each segment.

2. Number of items per Unit: The number of items to be developed for each CPT Unit will equal ten (10) times the number of segments making up the Unit.

3. Item difficulty: CPT items will possess high discriminative power; WLC will endeavor to construct items of relatively high initial difficulty and sensitivity to gain, using the general guideline that not more than 40% of the students should answer the item correctly on a pre-test, nor less than 50% of the students should answer the item correctly on a Post-Test. Actual indices of item difficulty will be determined in the validation testing program at the USNA, as a basis for subsequent item revision.

4. Format of items: CPT items will be developed in a final multiple-choice format; the student will be required to select the one correct or most appropriate answer from among four possible alternatives.

5. Types of items: CPTs will include items which represent "acquisition of knowledge" (Type I items) and "application of knowledge" (Type II items). Type I items will measure acquisition of specific factual information covered in one or more segments of a Unit. Acceptable forms of Type I items include definition-identification, discrimination-comparison, multiple discrimination, and concept identification questions.
   Type II items will measure application of concepts and principles covered in one or more segments of a Unit. Acceptable forms of Type II items include problem identification (selection of the correct or most

appropriate example or illustration of a given concept or principle) and problem solving (selection of the correct or most appropriate solution of a given problem). Each CPT will contain approximately equal numbers of Type I and Type II test items.

6. Length of items: CPT items will be concise; they will contain only essential details.
   As a specific guideline, each CPT item will be constructed so that a USNA midshipman may reasonably be expected to read, understand, and respond to the item within one minute.

B. Content Analysis and Item Construction

1. Analysis of content: The CPT should measure acquisition and application of content within segments, and also the ability to integrate and utilize material presented in several segments of the Unit.
   Construction of a CPT will therefore depend upon an analysis of Unit content.
   WLC will examine all terms, concepts, and principles contained in the Unit, together with their associated definitions, descriptions, purposes, advantages, disadvantages, and/or uses.
   Groups of terms, concepts, principles, etc., which have similar and/or potentially confusable elements will be identified. Test items can then be constructed which measure, in a multiple-choice format, the ability to identify, discriminate between, compare, and apply in problem situations, elements of the Unit content.
   Other relationships among terms, concepts, principles, etc., of the Unit will also be examined, so that test items can be constructed which measure, in a multiple-choice format, the ability to generalize from, integrate, and apply material presented in the Unit.

C. Classification of Items

1. Type I items (acquisition of knowledge)

   a. Definition-identification: Item will require selection of the correct definition, description, purpose, or use of a given term, concept, or principle;

      or:

      Item will require selection of the correct term, concept, or principle which is defined or described by a given definition or description.

   b. Discrimination-comparison: Item will require selection of the correct distinction between or comparison of a given set of terms, concepts, and/or principles;

      or:

Item will require the correct matching of a set of terms, concepts, and/or principles with a set of definitions and/or descriptions, as:

Which matching of words and statements is correct?

1. Term 1	A. Definition 1
2. Term 2	B. Definition 2
3. Term 3	C. Definition 3
4. Term 4	D. Definition 4

(possible answers, one of which is correct):

a) 1-B, 2-D, 3-B
b) 1-D, 2-B, 3-A
c) 1-D, 2-B, 3-C
d) 1-C, 2-D, 3-A

2. Type II items (application of knowledge)

a. Generalization-problem identification: Item will require selection of the correct or most appropriate "real-life" application, example or illustration of a given concept or principle;

or:

Item will require selection of the correct concept or principle illustrated by a given "real-life" example or illustration.

or:

Item will require correct matching of a set of concepts and/or principles with a set of "real-life" examples and/or illustrations.

b. Problem solving: Item will require selection of the correct or most appropriate solution of, resolution of, or reaction to a given "real-life" problem or situation;

or:

Item will require correct matching of a set of concepts and/or principles with a set of solutions of, resolutions of, and/or reaction to a given "real-life" problem or situation (i.e., how would alternative theories, methods, or approaches deal with the same situation or problem).

APPENDIX B

APPENDIX B

Experiment I

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Sequence	1	3.845	< 1	0.151	< 1	41.865	< 1	24.566	< 1
Frequency	1	21.701	< 1	10.682	< 1	899.486	3.11	16.239	< 1
Sequence X Frequency	1	0.938	< 1	213.767	< 1	157.827	< 1	133.232	1.35
Students in Groups	40	29.656		223.048		288.808		99.034	
Total Among Students	43								
Half	1	743.324	53.25**	4961.725	53.42**	6200.525	28.60**	2178.861	41.61**
Units in Half	2	229.795	26.76**	4398.961	50.42**	164.077	1.60	351.193	8.37**
Media	1	53.623	3.84	372.646	4.24*	512.719	2.36	211.355	4.04
Sequence X Unit in Half	2	19.001	2.21	107.974	1.23	286.230	2.78	3.357	< 1

Experiment I

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Frequency X Half	1	12.366	< 1	86.959	< 1	192.509	< 1	157.972	3.01
Frequency X Unit In Half	2	5.998	< 1	49.876	< 1	33.913	< 1	53.567	1.28
Media X Frequency	1	0.179	< 1	6.708	< 1	163.223	< 1	1.687	< 1
Sequence X Frequency X Unit In Half	2	32.032	3.73*	274.466	3.13*	19.194	< 1	74.088	1.77
Students X Half In Groups	40	13.959		87.942		216.794		52.363	
Students X Unit In Groups X Half	78[a]	8.588		87.593		102.826		41.952	
Total Within Students	130								

* $p < .05$

** $p < .01$

[a] Least squares estimation of two missing values required the subtraction of two degrees of freedom from this source.

APPENDIX C

APPENDIX C

Experiment I

Means for Unit - Method Combinations

Total CPT Scores

Group	Frequency	Sequence 1			
		Videotape		Audiotape	
		Unit 1	Unit 3	Unit 7	Unit 9
1	HRDF	16.83	10.42	17.01	19.50
2	LRDF	13.36	11.00	17.91	18.09
		Sequence 2			
		Audiotape		Videotape	
		Unit 1	Unit 3	Unit 7	Unit 9
3	HRDF	16.00	12.45	17.36	16.18
4	LRDF	16.20	10.20	17.36	16.00

Type I Percentages

Group	Frequency	Sequence 1			
		Videotape		Audiotape	
		Unit 1	Unit 3	Unit 7	Unit 9
1	HRDF	71.02	46.25	68.85	73.50
2	LRDF	61.65	48.18	69.84	69.14
		Sequence 2			
		Audiotape		Videotape	
		Unit 1	Unit 3	Unit 7	Unit 9
3	HRDF	67.99	51.36	66.94	64.26
4	LRDF	72.18	47.00	71.82	66.41

Experiment I

Means for Unit - Method Combinations

Type II Percentages

Group	Frequency	Sequence 1			
		Videotape		Audiotape	
		Unit 1	Unit 3	Unit 7	Unit 9
1	HRDF	61.27	55.83	65.46	74.32
2	LRDF	55.62	48.18	69.70	72.74
		Sequence 2			
		Audiotape		Videotape	
		Unit 1	Unit 3	Unit 7	Unit 9
3	HRDF	60.96	60.91	70.20	68.27
4	LRDF	54.70	54.00	64.30	61.67

Progress Check Percentages

Group	Frequency	Sequence 1			
		Videotape		Audiotape	
		Unit 1	Unit 3	Unit 7	Unit 9
1	HRDF	66.67	77.42	83.42	76.57
2	LRDF	65.45	67.85	75.76	76.12
		Sequence 2			
		Audiotape		Videotape	
		Unit 1	Unit 3	Unit 7	Unit 9
3	HRDF	66.36	73.74	78.55	75.45
4	LRDF	72.75	73.81	77.61	74.47

APPENDIX D

APPENDIX D

Experiment II

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Frequency	1	60.004	4.25*	593.408	3.17	708.701	3.08	242.592	1.88
Sequences	2	6.364	< 1	179.668	< 1	29.806	< 1	28.847	< 1
Frequency X Sequence	2	3.873	< 1	130.748	< 1	33.591	< 1	84.341	< 1
Students in Groups	38	14.128		187.186		229.729		128.988	
Total Among Students	43								
Units	2	182.829	40.77**	4466.016	40.24**	1825.830	17.17**	3817.990	65.39**
Method	2	2.307	< 1	73.624	< 1	1.862	< 1	225.302	3.86*
Frequency X Units	2	2.573	< 1	37.826	< 1	113.469	1.06	4.472	< 1
Frequency X Method	2	1.749	< 1	49.096	< 1	52.358	< 1	16.238	< 1

Experiment II

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Unit X Method	2	0.277	< 1	8.568	< 1	20.968	< 1	13.183	< 1
Frequency X Units X Method	2	3.070	< 1	126.534	1.14	41.615	< 1	13.229	< 1
Students X Unit-Method Combinations in Groups	75[a]	4.484		110.971		106.351		58.389	
Total Within Students	87								

* $p < .05$

** $p < .01$

[a] Least squares estimation of a missing value required the subtraction of one degree of freedom from this source.

APPENDIX E

APPENDIX E

Experiment II

Means for Unit - Method Combinations

Total CPT Scores

Group	Frequency	Unit 4	Unit 6	Unit 10
		Overt-written	Overt-spoken	Covert
1	HRDF	20.86	21.87	17.71
4	LRDF	19.57	20.43	17.41
		Overt-spoken	Covert	Overt-written
2	HRDF	21.86	21.86	18.00
5	LRDF	19.38	20.50	15.75
		Covert	Overt-written	Overt-spoken
3	HRDF	20.43	21.14	16.57
6	LRDF	19.12	19.25	17.00

Experiment II

Means for Unit - Method Combinations

Type I Percentages

Group	Frequency	Unit 4	Unit 6	Unit 10
		Overt-written	Overt-spoken	Covert
1	HRDF	68.36	85.16	72.26
4	LRDF	67.33	77.67	58.16
		Overt-spoken	Covert	Overt-written
2	HRDF	73.46	82.14	61.21
5	LRDF	62.50	78.11	53.56
		Covert	Overt-written	Overt-spoken
3	HRDF	63.24	78.57	57.13
6	LRDF	66.96	71.10	58.91

Experiment II

Means for Unit - Method Combinations

Type II Percentages

Group	Frequency	Unit 4	Unit 6	Unit 10
		Overt-written	Overt-spoken	Covert
1	HRDF	70.51	59.04	55.36
4	LRDF	63.37	57.13	56.27
		Overt-spoken	Covert	Overt-written
2	HRDF	72.31	62.24	58.94
5	LRDF	66.40	57.15	51.55
		Covert	Overt-written	Overt-spoken
3	HRDF	72.31	61.23	53.59
6	LRDF	60.96	56.22	54.69

Experiment II

Means for Unit - Method Combinations

Progress Check Percentages

Group	Frequency	Unit 4	Unit 6	Unit 10
		Overt-written	Overt-spoken	Covert
1	HRDF	81.80	66.26	83.87
4	LRDF	79.83	61.74	82.93
		Overt-spoken	Covert	Overt-written
2	HRDF	77.50	62.24	85.71
5	LRDF	79.83	61.74	82.93
		Covert	Overt-written	Overt-spoken
3	HRDF	81.79	72.94	83.86
6	LRDF	76.72	65.64	79.42

APPENDIX F

APPENDIX F

Experiment III

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Method	3	14.692	1.70	35.49	< 1	316.14	3.40*	55.29	1.67
Students in Method	27[a]	8.664		109.36		92.83		33.03	
Units	1	25.611	3.18	198.38	2.92	1728.34	17.34**	480.42	22.94**
Method X Unit	3	13.324	1.66	176.08	2.60	40.91	< 1	10.36	< 1
Students X Unit in Method	27[a]	8.045		67.85		99.69		20.94	

* $p < .05$

** $p < .01$

a df = 30 in the analysis of Progress Check data.

APPENDIX G

APPENDIX G

Experiment III

Means for Media - Method Combinations

Total CPT Scores

	HRDF-HMF	HRDF-MMF	HRDF-LMF	LRDF-LMF
AT-IP	25.00	25.57	26.00	25.37
CAI	25.60	27.28	26.91	29.50

Type I Percentage

	HRDF-HMF	HRDF-MMF	HRDF-LMF	LRDF-LMF
AT-IP	69.52	70.06	67.55	64.88
CAI	62.00	62.14	61.36	71.62

Type II Percentage

	HRDF-HMF	HRDF-MMF	HRDF-LMF	LRDF-LMF
AT-IP	54.74	67.69	62.21	61.83
CAI	66.00	74.29	73.18	76.88

Progress Check Percentage

	HRDF-HMF	HRDF-MMF	HRDF-LMF	LRDF-LMF
AT-IP	83.77	81.77	79.27	80.76
CAI	89.72	86.31	86.72	84.60

APPENDIX H

APPENDIX H

Experiment IV

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Squares	1	4.182	< 1	1.065	< 1	267.614	1.01	.001	< 1
Sequences In Squares	4	6.766	< 1	127.017	< 1	159.428	< 1	2.319	< 1
Students In Groups	38	20.258		313.217		265.213		85.577	
Total Among Students	43								
CPT Units	2	161.938	18.74**	1120.092	6.02**	3351.945	24.46**	1551.320	24.56**
Remediation Method	2	8.346	< 1	316.186	1.70	93.775	< 1	95.275	1.51
Squares X Units	2	12.307	1.42	30.808	< 1	541.992	3.96*	50.073	< 1
Squares X Method	2	2.798	< 1	55.034	< 1	152.459	1.11	58.489	< 1
Unit X Method In Squares	4	5.592	< 1	239.461	1.29	26.850	< 1	98.763	1.56

Experiment IV

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Students X Unit-Method Combinations in Groups	75[a]	8.641		186.082		137.023		63.168	
Total Within Students	87								

* $p < .05$

** $p < .01$

[a] Estimation of missing value required the subtraction of one degree of freedom from this source.

APPENDIX I

APPENDIX I

Experiment IV

Means for Unit - Method Combinations

Total CPT Scores

Sequence	Square 1			Square 2		
	Unit 2	Unit 8	Unit 12	Unit 2	Unit 8	Unit 12
1	High	Low	None	High	None	Low
	20.00	16.14	18.14	19.43	14.29	17.71
2	Low	None	High	Low	High	None
	19.50	17.25	17.75	22.00	16.14	17.86
3	None	High	Low	None	Low	High
	19.00	16.57	19.78	19.62	16.12	17.75

Type I Percentages

Sequence	Square 1			Square 2		
	Unit 2	Unit 8	Unit 12	Unit 2	Unit 8	Unit 12
1	High	Low	None	High	None	Low
	68.09	57.99	59.04	67.23	56.31	63.83
2	Low	None	High	Low	High	None
	70.60	63.98	55.00	75.64	63.03	61.89
3	None	High	Low	None	Low	High
	65.54	57.14	72.91	67.66	58.08	58.24

Experiment IV

Means for Unit - Method Combinations

Type II Percentages

Sequence	Square 1			Square 2		
	Unit 2	Unit 8	Unit 12	Unit 2	Unit 8	Unit 12
1	High	Low	None	High	None	Low
	61.39	49.46	61.90	61.51	35.19	54.27
2	Low	None	High	Low	High	None
	57.68	49.05	63.34	70.30	41.77	57.14
3	None	High	Low	None	Low	High
	60.41	52.76	61.30	62.48	48.10	60.85

Progress Check Percentages

Sequence	Square 1			Square 2		
	Unit 2	Unit 8	Unit 12	Unit 2	Unit 8	Unit 12
1	High	Low	None	High	None	Low
	78.59	83.94	79.89	73.33	80.80	88.33
2	Low	None	High	Low	High	None
	76.25	84.78	83.54	72.40	84.81	86.37
3	None	High	Low	None	Low	High
	70.00	85.71	87.41	75.01	85.55	83.54

APPENDIX J

APPENDIX J

Experiment V

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	F	Type I Mean Square	F	Type II Mean Square	F	Progress Check Mean Square	F
Groups	1	16.43	-1.70	12.39	3.41	0.29	< 1	8.38	1.28
Students in Groups	40	9.64		3.63		3.58		6.54	
Total	41	402.00							

APPENDIX K

APPENDIX K

Media Differences

Analysis of Variance Based on Unweighted Means

Sources of Variance	Degrees of Freedom	Total CPT Mean Square	Total CPT F	Type I Mean Square	Type I F	Type II Mean Square	Type II F	Progress Check Mean Square	Progress Check F
Among Media	5	16.114	< 1	25.263	< 1	45.313	< 1	103.429	2.44
(Repeated Media)	(2)	(21.026)	(<1)	(22.847)	(< 1)	(35.113)	(<1)	(38.316)	(<1)
(Remainder)	(3)	(12.840)	(<1)	(26.874)	(< 1)	(52.112)	(<1)	(146.837)	(3.46)
Units in Media	7	49.489		81.256		55.035		42.441	
Total	12								

APPENDIX L

APPENDIX L

Unweighted Mean Percent Correct for Media and Cumulative Posttest Units

CPT Unit	Segments	Total CPT	Type I	Type II	Progress Check
		Experiment One			
	Audiotape Lecture with Panel Booklet (AT-PB)				
1	2.2 - 2.5	63.99	68.21	58.14	69.56
3	3.1 - 3.4	52.54	48.20	54.73	73.78
7	5.7 - 5.10	68.52	69.36	67.42	79.59
9	7.1 - 7.4	68.61	68.33	69.25	75.59
	Average	63.42	63.52	62.38	74.63
		Experiment Two			
	Linear Text (LT)				
4	4.1 - 4.3	67.34	66.98	67.64	78.92
6	5.4 - 5.6	69.47	78.79	58.84	65.52
10	8.1 - 8.3	56.76	58.71	55.07	83.48
	Average	64.52	68.16	60.52	75.97

Unweighted Mean Percent Correct for Media

and Cumulative Posttest Units

CPT Unit	Segments	Total CPT	Type I	Type II	Progress Check
		Experiment Three			
	Audiotape and Instrinsically Programmed Booklet (AT-IPB)				
5	4.4 - 4.7	64.96	68.00	61.62	81.39
	Computer Assisted Instruction (CAI)				
13	12.1 - 12.2	68.31	64.28	72.59	86.84
		Experiment Four			
	Syndactic Text (ST)				
2	2.6 - 2.8	66.42	69.13	62.30	74.26
8	6.1 - 6.3	53.62	59.42	46.06	84.26
12	11.1 - 11.3	60.55	61.82	59.80	84.85
	Average	60.20	63.46	56.05	81.12
		Experiment Five			
	Learning Activities Summary (LAS)				
11	8.4 - 8.6	66.37	73.88	57.79	77.32
	Unweighted Media Average	64.63	66.88	61.82	79.54