HARLESS, William Glenn, 1933-
THE USE OF COMPUTER-ASSISTED INSTRUCTION TO TEST THE TOTAL TIME HYPOTHESIS IN VERBAL CONCEPT LEARNING.

The University of Oklahoma, Ph.D., 1969
Education, theory and practice

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

# THE USE OF COMPUTER-ASSISTED INSTRUCTION TO TEST THE TOTAL TIME HYPOTHESIS IN VERBAL CONCEPT LEARNING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

WILLIAM GLENN HARLESS

Oklahoma City, Oklahoma

1968

THE USE OF COMPUTER-ASSISTED INSTRUCTION TO TEST THE

TOTAL TIME HYPOTHESIS IN VERBAL CONCEPT LEARNING

APPROVED BY

DISSERTATION COMMITTEE

## ACKNOWLEDGMENTS

Sincere thanks are extended to Dr. Harry J. Parker for advice and assistance in the development of this paper and to Dr. Edward N. Brandt, Jr. for his guidance throughout my graduate program. I am also grateful for the suggestions of Doctors W. W. Schottstaedt, T. N. Lynn, G. D. Adams, and J. G. Bruhn, who served on the reading committee.

My appreciation is also extended to Dr. Arthur W. Nunnery, who willingly provided the subjects and subject matter for this experiment, Dr. J. N. Kanak, who was instrumental in the formulation of the research, and to Dr. R. C. Duncan for his advice in the analysis of the data. And finally, I am indebted to Mrs. Nancy Lucas for her competent help in all phases of the experiment and for preparing the manuscript.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF ILLUSTRATIONS

# THE USE OF COMPUTER-ASSISTED INSTRUCTION TO TEST THE TOTAL TIME HYPOTHESIS IN VERBAL CONCEPT LEARNING

## CHAPTER I

## INTRODUCTION

### The Role of Computer-Assisted Instruction in a Verbal Learning Experiment

The use of computers in the educational process, while still in the embryonic stage, shows exceptional potential. The state of computer science is rapidly advancing to the point where there will be no practical limitations on computer storage capacities, speeds, and transmission capabilities. The ability of computers to service several users at remote stations (terminals) on a time-shared basis permits its use as an aid in information retrieval and dissemination. One such use currently receiving a great deal of attention is computer-assisted instruction, which may be defined as the utilization of computers to disseminate course information and other material relevant to the formal education of the student. This technique is known as computer-assisted instruction (CAI), computer-aided instruction, or computer-assisted learning; however, the first seems to be the most popular term.

It is generally agreed that pedagogical approaches in the form of computer-assisted instruction programs can be categorized into four broad classes: (1) drill; (2) tutorial; (3) simulation; and (4) problem solving (Bunderson, 1966).

The drill program is limited primarily to the development of vocabulary skills and usually consists of a rigid format for student-computer interaction. It is strictly a method of reinforcement and has evolved naturally into the repertoire of computer-assisted instruction from "programmed instruction" and the "teaching machine," both of which were designed for this type of instruction.

Simulation programs have been developed to permit the student to project his knowledge into a simulated situation in order to learn to make decisions and develop skills which are concerned with that situation. Such a learning program in the medical environment was developed by Feurzeig (1964) and dealt with the diagnostic process. Similar work was done by Entwisle (1963) and involved the computerized simulation of a patient. This patient was queried by the student. On the basis of the information obtained, the student selected a diagnosis from a list of six possible diseases. If his diagnosis was inaccurate, the computer suggested further questions. The primary difference between the two studies was that certain pedagogical additives were present in Feurzeig's study that required the student to obtain the information in a logical fashion.

The problem-solving approach involves the use of the computer as an independent resource for mathematical and statistical investigation. For example, a student in Biostatistics might be required to design an experiment. A library of statistical routines stored in the computer would be readily available for use in the solution of his problem. In this approach the computer would, in effect, be used as a sophisticated, powerful calculator to give the student the ability to either create his own algorithms via the computer terminal or to use those already stored in the data bank.

The tutorial approach is used as a method of presenting independent resource material, primary material, and reinforcement material. It is basically an interactive session where the computer is programmed to present information employing various pedagogical strategies.

The use of the computer as the instructional medium for these four approaches indicates attributes which are not found in traditional programmed instruction techniques. These are: (1) an advanced interface providing dynamic presentation capabilities; (2) a natural language environment; (3) unlimited branching capability; and (4) an automatic recording of student responses.

The most advanced terminal interface currently available is a cathode ray tube (CRT), which is similar to a television screen. Connected to it are a standard typewriter keyboard with several special characters and a light pen which contains a photoelectric cell capable of sensing light elicited by the screen. The student is therefore provided two means of interacting with the computer: typing on the keyboard and pointing with the light pen.

The ability of the computer to analyze a natural language response from the terminal keyboard allows the student to function in a more natural environment with the feeling of conversational interaction. This freedom of communication enhances the flexibility of both the interactive session and the evaluation process. Traditional programmed instruction texts make the student interact with the material through multiple-choice questions. The evaluation of the student's learning is, therefore, restricted to the recognition technique of multiple-choice questions. This may be analogous to an objective examination which measures the student's ability to recognize the information in context. The computer's ability to accept natural language responses is analogous to the subjective examination which ideally measures the student's ability to freely recall information and apply it. Obviously, this is a more difficult task for the student but is potentially a more meaningful measure of learning.

The branching power of the computer has no significant limitations. In course development this is an extremely powerful attribute that carries with it some rather imposing responsibilities concerning the quality and flexibility of the interactive situation. The instructor is given the ability to control the learning situation at virtually any level he chooses. Once he determines the criterion for learning, he may branch the student to the appropriate level of remedial or advanced information depending on the student's relative mastery of the material.

The computer's record-keeping ability is perhaps the most important attribute it brings to the educational task. Everything the student does while participating in a didactic lesson at a terminal is recorded on some storage device; i.e., magnetic tape or disk. Embodied in this record is an identifying number, the student's response, and other information concerning the student's progress at the time the record was made. Data-analysis programs can be written to process this information and present it in either raw or summary form. This provides immediate feedback to the instructor on the pedagogical effectiveness of the presentation. This attribute, coupled with the natural language environment, also provides the investigator with data concerning normal student behavior during the acquisition of information.

The physical attributes of the computer are, therefore, quite clear; however, the most effective use of these attributes in the development of a learning situation is not so well defined. While educational interest in the possibilities of a computer-based instructional system is high, experience has shown that the successful operation of such a system presents some unique problems which cannot be solved by traditional computer techniques. These problems revolve around the general subject of efficient and effective pedagogy. One must determine how to efficiently use the student's time while he is engaged in the interactive session with the computer, but efficiency must in no way compromise the effectiveness of the learning situation. Efficiency here means the use of the student's time, and effectiveness means the student's retention and behavioral changes. Both of these depend on the instructor's ability to determine the appropriate levels of mediation and control that are provided by the computer in the interactive session. For the purposes of this paper, mediation is defined as the process of associating information to be learned to information already known. The mediators resulting from this process act as aids to memory.

Some insight into these problems has been gained from research in the verbal learning laboratory. Research in verbal learning has been abundant and is concerned with the acquisition, transfer, and retention of such materials as paired-associate, serial, and free recall lists of nonsense syllables. Some of the literature concerning these three experimental tasks is reviewed in the next chapter.

In particular, the Bugelski total time hypothesis was developed in the verbal learning laboratory, and it has definite implications in practical learning situations. Succinctly, this hypothesis is that the total time to learn a list of materials presented at varying rates of exposure to different groups of subjects is a constant (Bugelski, 1962). That is, even though the total number of trials to learn a set of materials may differ as a function of the exposure time for the materials, the total time required to learn these materials may be the same across different rates of exposure. The evidence indicates that the hypothesis can be expected to hold whenever task requirements do not exceed simple rehearsal and whenever effective time (the time during which the subject is attending to the task) bears a positive linear relationship to nominal time (the total amount of exposure time for each item). This finding suggests that in a practical situation the learning process is dependent upon the total exposure time of the material and is not affected by the number of exposures.

Since the total time hypothesis had not been systematically investigated within a practical learning environment, it seemed worthwhile to relate this finding in the verbal learning laboratory to the practical learning situation. Such research provided insight into the problem of pedagogy for CAI which, in turn, provided a unique and desirable setting for such research to be accomplished. If acquisition of conceptual material could be shown to be dependent upon exposure time, then CAI presentations could be developed to pace the student through a lesson at the most effective learning rate. The absence of this relationship would indicate individual differences, and presentations could be self-paced in consideration of these differences. The aforementioned attributes of CAI, advanced interface, natural language analysis, unlimited branching power, and detailed record keeping, provided unprecedented power in the execution of this verbal learning study in the practical learning environment. By using information germane to the lesson being presented, the pupil was able to fulfill both roles of student and subject simultaneously. It was felt that this study could act as an indicator of the relationship between the findings of the abundant nonsense-syllable learning studies and higher cognitive processes. The establishment of such a relationship would put learning theorists far ahead in their search for effective learning methodology.

## The Statement of the Problem

It was therefore the purpose of the study to investigate the implications of the Bugelski total time hypothesis at the verbal concept level. This included rates of acquisition and measures of retention across various experimental conditions. The IBM 1500 instructional system was used to collect other meaningful data concerning the subjects' ability to use the terminal interface and the computer's ability to pace the subject through a learning situation. Specifically, this study tested the following hypothesis: There is a fixed amount of time required to learn a fixed amount of conceptual material, and this time is independent of the number of exposures of the information.

The following sub-hypotheses were also tested: (1) Retention of the experimental material is independent of the rate of presentation; (2) The time used to recall the experimental material is independent of the rate of presentation; and (3) Total time attending to the task is independent of the rate of presentation.

## CHAPTER II

## A REVIEW OF THE LITERATURE

### The Verbal Learning Experiment

Contemporary research in the verbal learning laboratory primarily involves the analysis of three tasks: (1) paired-associate lists; (2) serial learning lists; and (3) free recall lists.

The general method employed in paired-associate learning studies is as follows: The subject is presented a specified number of paired trigrams which are usually three-letter nonsense syllables. Figure 1 is an example of such a list. The trigrams on the left in Figure 1 are the stimuli, and the juxtaposed items on the right are the responses. It is the responsibility of the subject to make the association between these trigrams during the study (acquisition) phase of the experiment. In the testing (recall) phase of the experiment, the subject is required to recall the appropriate response when shown the stimulus.

| | |
|---|---|
| BAF | JOC |
| DEX | SIV |
| COL | WEY |
| LIG | HOK |
| CAG | MUZ |

Fig. 1.--A paired-associate list of nonsense trigrams.

Serial learning studies consist of the sequential acquisition of a set of items (usually nonsense trigrams). During the acquisition phase the items are exposed individually via a standard device (memory drum, slide projector, etc.) at a controlled rate. In the recall phase the subject is required to provide each of these items in its serial sequence.

Free recall learning provides the subject with a list of trigrams in the acquisition phase. In the recall phase he is required to remember as many items in the list as possible. Unlike the serial task, the free recall task does not require that the subject recall the items in a sequential order. It is also differentiated from the paired-associate task by the absence of any stimulus item in the recall phase.

## The Bugelski Total Time Hypothesis

This particular study was the result of a finding in 1962 by B. R. Bugelski of the University of Buffalo. Bugelski's verbal learning experiment concerned the rate of presentation, total presentation time, and mediation in paired-associate learning. His hypothesis was stated as follows:

> It is the present hypothesis that in a least some areas of memorization, and under some conditions of presentation, the degree of learning will be a function of total time, regardless of the duration of the individual trials or interitem times (Bugelski, 1962).

Bugelski's experiment was actually designed to test the claims of Rock (1957) concerning his one-trial learning experiment. The Bugelski method was as follows: The Hunter card master was used to expose eight pairs of nonsense syllables used in the Rock experiment at varying presentation times. The stimulus syllable was exposed for two seconds and remained visible while the response syllable was exposed for either 2, 4, 6, 8, or 15 additional seconds. Two seconds elapsed before a new stimulus syllable appeared. It should be clear that the total time available to the subject for the acquisition of a given pair varied from 6 seconds for the 2-second response group to 19 seconds for the 15-second response group. The 6 seconds for the first group included 2 seconds for the stimulus exposure, 2 seconds for the stimulus and response together, and 2 seconds between trials.

Each subject learned the same eight pairs of syllables to a criterion of two successful anticipations of the complete list. The apparatus did not allow for the elimination of pairs as they were learned. This was a mechanical limitation of the experimental equipment and a possible source of confounding due to the facilitation of learning by serial cues and contextual cues within the list (Deese and Kaufman, 1957).

The findings of the experiment supported the original hypothesis that total learning time was a significant variable to consider in at least some kinds of learning. Bugelski concluded that Rock's design was inappropriate for measuring the one-trial learning phenomenon due to the excessive time allowed for the trial.

## The Related Literature

The implications of the Bugelski experiment were much broader than simply an argument against the Rock one-trial experiment. The suggestion that learning processes existed which were a function of the time exposure to material stimulated Bugelski and Rickwood (1963) to replicate the previous Bugelski experiment allowing a group of subjects to control their own exposure times. The mean total exposure time for this self-pacing group was not significantly different from the mean of the five groups of the original Bugelski study. This experiment supported the existence of a total time phenomenon.

The validity of the Bugelski total time hypothesis has been investigated in all three of the verbal learning tasks discussed earlier. The total time hypothesis has enjoyed almost universal support from the work done in paired-associate and free recall lists. In serial learning, experiments testing the hypothesis are less numerous and their results are less conclusive.

In general, two procedures are used to test the total time hypothesis. In one, total learning time is held constant while different groups of subjects are presented the to-be-learned material at different rates. The number of items correct is compared under the various conditions. For example, the performance on the first trial of a group learning at a four-second rate is compared with the performance on the second trial of a group learning at a two-second rate. The measure is the number of correct items on the recall phase. In the second procedure, learning is carried to the same criterion for different groups of subjects presented the to-be-learned material at different rates. The measure for this procedure is the total time to reach criterion.

Recent research shows that the validity of the hypothesis seems to depend upon compliance to the following conditions: (1) The task does not exceed simple rehearsal; (2) Effective time bears a positive relationship to nominal time; and (3) The motivational and perceptual thresholds of the subjects are not compromised.

The evidence indicates that the total time relationship holds in tasks which require only simple rehearsal for mastery. In one experiment, Glucksberg and Laughery (1965) categorized the experimental task on the basis of operations that a subject must perform to master the task. This study provided a criterion that differentiated the tasks for which the total time hypothesis was valid. Paired-associate and free recall lists did not appear to require operations other than rehearsal or study time. On the other hand, the nature of complicated serial learning tasks such as the Glucksberg-Laughery experiment required active mental processes more involved than rote memorization.

The literature also shows that a distinction must be made between the time potentially available for learning and the time during which learning is actually occurring. This is classically called the difference between nominal and effective time (Kausler, 1966, p. 259). Nominal time may be defined as clock time, the time potentially available for repeated rehearsal. Effective time is defined as that part of the nominal time during which repeated rehearsals are actually evoked. Research by Carroll and Burke (1965) and by Nodine (1965) suggests that increases in the stimulus-response presentation time decreases the number of trials required to reach criterion. However, an increase in stimulus time alone fails to produce this effect. This implies that the increased nominal (stimulus) time is not all used as effective (rehearsal) time and, therefore, that learning does not bear a one-to-one relationship to exposure time. This finding is contradictory to the Bugelski hypothesis.

Extreme presentation rates have also led to the breakdown of the relationship between total time and learning. Johnson (1964) presented eight items in a paired-associate list in which the stimulus was a consonant-vowel-consonant (CVC) trigram and the response was a digit between 1 and 8. The design of the experiment was a 4 x 4 factorial. The time to learn per item was set into either 10, 20, 40, or 80 seconds. Each of these times was divided into either 1, 5, 10, or 20 exposures, depending on the group with which the subject was involved. An analysis of variance indicated that with total time of exposure held constant, the frequency of exposure did have a significant effect. The Johnson experiment provided evidence that extremely slow rates led to ineffective use of time for study and extremely fast rates tended to inhibit the acquisition of correct responses. The research seemed to indicate upper and lower limits of presentation time beyond which optimal study did not take place. These limits were defined respectively as the "motivational threshold" and the "perceptual threshold."

In general, the Bugelski total time hypothesis has found support in verbal learning experiments. The following nine paired-associate learning studies support it unconditionally: Baumeister and Hawkins (1966); Bugelski (1962); Bugelski and Rickwood (1963); Goss, Morgan, and Golin (1959); Hovland (1949); Newman (1964); Postman and Goggin (1966); Underwood and Keppel (1963); and Wilcoxon, Wilson, and Wise (1961). Three free recall learning studies (Murdock, 1965; Bousfield, Sedgewick, and Cohen, 1954; and Deese, 1957) and one serial list study (Keppel and Rehula, 1965) support the total time hypothesis with no reported contradictory results. The Carroll and Burke, Nodine, and Johnson experiments discussed earlier indicate conditions of the paired-associate task which must be present for the hypothesis to be valid. Also, three serial list studies (Postman and Goggin, 1964; Glucksberg and Laughery, 1965; and Braun and Heymann, 1958) present evidence contraindicating its validity.

Other research has been accomplished that is equally as important as the work with the total time hypothesis in the development of computer-assisted instruction materials. Such matters as the meaningfulness-acquisition relationship (Noble, 1952), mediation and mediated transfer (Peterson, Culaveta, and Sheahan, 1964), nominal and functional stimuli

(Underwood, 1963), and interference theory with reference to retroactive and proactive inhibition (Ceraso, 1967) have been studied. These studies have had a directional effect on the investigator in the design of this experiment and the inferences made from the data. The influence of each will become apparent at various times throughout this study.

# CHAPTER III

# METHOD

## Subjects and Experimental Material

The subjects used in this study were medical students at the University of Oklahoma School of Medicine who, as a regular part of their training, were studying infant nutrition in pediatrics. These subjects were involved in the traditional medical school program in which they progressed as a body through the first two basic science years and then separated into groups for the last two years of clinical training in the School's various disciplines. It was during their third and fourth years that this group rotated through the pediatric service. Forty juniors and twelve seniors were involved with the pediatric rotation during the term of the experiment (December, 1967, to July, 1968).

These students comprised the sample of the present study and were considered equally competent to acquire the experimental material. Therefore, they were treated as a homogeneous group. This assumption was based on consultation with the instructor of pediatrics, who stated that the subjects had not been formally exposed to these concepts at any time during their previous training.

The experiment was embodied in a computerized tutorial session and consisted of material that was germane to the subjects' pediatric service training. Specifically, material on infant nutrition, ordinarily presented by lecture, was presented to the subjects via computer-assisted instruction techniques. The experimental material, which consisted of 15 different facts concerning vitamin functions and requirements, was incorporated at the beginning of the tutorial session. The subject material on infant nutrition which followed the experimental material was livelier and more directly related to the subjects' pediatric training. It was felt that placing the subject matter after the experimental material would minimize the likelihood of outside discussions among the subjects concerning the experiment.

## Treatment Groups

The subjects in the experiment were randomly assigned to four treatment groups, each of which was to view the experimental material at a different time exposure rate. These rates were established from empirical evidence provided by individuals unassociated with the study. Specifically, twelve individuals whose backgrounds ranged from computer programmers to graduate school faculty judged the amount of time required to read each fact as it appeared on the CRT. As each fact was presented, each judge read it and pressed the space bar. This time was recorded by the computer. Subsequently, these response times were retrieved from the system and averages for each question computed. These averages became the pre-experimental reading times for each fact to be reviewed by the CAI staff. The staff was made cognizant of the importance of finding the perceptual threshold for each fact since this was one of the conditions of the total time hypothesis. Too little time for any fact would compromise the perceptual threshold of many subjects and, obviously, have a debilitating effect on his acquisition of that fact. Too much time would provide subjects with extra rehearsal time and facilitate their learning of the item. In either case, confounding would occur.

One refinement was added to assure realistic exposure times with respect to the degradation problem that results from multiple users on the 1500 system. This investigator reviewed the experimental material with two terminals operating simultaneously since the subjects were expected to come in pairs for the lesson. As the investigator vocalized each item appearing on the screen, the staff made notes of those items with excessive or inadequate exposure times. Using this information, appropriate adjustments were made. Table 1 shows the pre-experimental reading times, final exposure times, and the experimental study items.

Each of the four treatment groups was presented the study material at a different time exposure rate. The first group was the "minimum time group" (M). For these subjects, each fact was presented for the duration of its determined perceptual threshold as discussed above. The second group, called the "comfortable time group" (C), was allowed to see each fact twice as long as group M. The third group was the "excessive time group" (E). They saw each fact four times as long as group M. The fourth group, the "self-paced group" (SP), acted as a control. For this group a fact stayed on the screen until the subject indicated he was ready to see the next item by pressing the space bar.

## Orientation and Instructions to Learn

Before the subjects came to the CAI laboratory, they were given a list of instructions explaining the following uses of the terminal: (1) how to initiate the session; (2) how to enter a response to the computer; (3) how to edit a response before entering it; and (4) how to

TABLE 1

EXPERIMENTAL MATERIAL AND ASSOCIATED MINIMUM EXPOSURE TIMES*

| Item Number | Pre-Experiment Mean Reading Time (Seconds) | Final Exposure Time (Seconds) | Experimental Items |
|---|---|---|---|
| 1 | 3.6 | 1.5 | One Vitamin A deficiency is nyctalopia (night blindness). |
| 2 | 5.0 | 3.0 | A disease caused by excess amounts of Vitamin A concentrate is lipemia. |
| 3 | 5.5 | 3.0 | Cocarboxylase is a coenzyme formed when thiamine combines with phosphates. |
| 4 | 4.6 | 2.6 | The daily requirement of thiamine is 1 to 4 mg., according to size. |
| 5 | 4.7 | 3.2 | Riboflavin is the vitamin which is responsible for the respiratory enzyme system. |
| 6 | 5.0 | 2.5 | One symptom of riboflavin deficiency is vascularization of the cornea. |
| 7 | 6.2 | 6.2 | The coenzyme I of niacin is known as diphosphopyridine nucleotide, or coxymase. |
| 8 | 4.9 | 4.9 | Niacin's coenzyme II is tri-phosphopyridine, or coferment. |

TABLE 1--Continued

| Item Number | Pre-Experiment Mean Reading Time (Seconds) | Final Exposure Time (Seconds) | Experimental Items |
|---|---|---|---|
| 9 | 4.1 | 3.1 | Pellagra is the disease caused by a niacin deficiency. |
| 10 | 3.4 | 3.4 | The infant requirement of Vitamin C is 30 mg. |
| 11 | 4.2 | 4.2 | Breast milk contains 4-7 mg. of Vitamin C per 100 cc. |
| 12 | 3.8 | 3.8 | Vitamin D is responsible for the metabolism of phosphorus and calcium. |
| 13 | 4.6 | 4.6 | Fish liver oils, irradiated foods, and sunshine are good sources of Vitamin D. |
| 14 | 3.6 | 3.6 | Vitamin K is an essential part of prothrombin formation. |
| 15 | 2.5 | 3.0 | Chlorophyll is one source of Vitamin K. |

* Times are for minimum time group. The comfortable time group was exposed to the facts twice as long and the excessive time group four times as long as the minimum time group.

terminate the session. The list also provided each subject with his identifying number for the lesson. It was made clear that the CAI session was not an examination, but a technique of self-study. The subjects were aware that their responses were being recorded by the computer in order to study the effectiveness of this technique. They were not told of the specific experiment nor that they were to fulfill any role other than that of student.

Each subject received individual attention from a member of the CAI staff when he came to the computer facility. This consisted of helping the subject interpret his list of instructions and getting him started on the lesson. When the subject initiated the session at the terminal, the computer led him through a brief orientation course on its own use. This was primarily a short, light, interactive session concerning the mechanics of entering a keyboard response and using the light pen.

This short session served two functions. First, it acted as a warmup session for the learning experiment; that is, the development of a set (e.g., postural adjustments, etc.) which maximized the subject's performance proficiency (Kausler, 1966, pp. 360-361). Second, it provided a measure of the subject's ability to function with the equipment. The acquisition of this measure was facilitated by the computer's timing capability. The subjects were requested by the computer to type the words <u>idiopathic</u>, then <u>thrombocytopenia</u>, and finally their telephone numbers. Each of these tasks was timed by the computer and their sum defined as the subject's psychomotor index. Figure 2 shows an example of an interactive session between the subject and the computer to determine the psychomotor index.

After the orientation session, the instructions for the learning

Computer: To ENTER a response:

1. While holding down the ALTERNATE CODE key---located in the upper left hand corner of the keyboard

2. Depress the SPACE BAR

This two-step procedure will cause the K to vanish and the response to be analyzed by the system.

Student: (Presses the space bar)

Computer: Ok, now it's your turn. Type your name and enter it by this procedure.

Student: John Doe

Computer: You made it, John Doe.

Let's practice that for a minute...
Idiopathic means a disease of unknown origin.
Type idiopathic.

Student: Idiopathic

Computer: Good. Now type thrombocytopenia.

Student: Thrombocytopenia

Computer: Not bad. And last, type your telephone number.

Student: 236-1366

Computer: Thanks. (You may call me at 525-7571.)

Fig. 2.--Interactive session to measure psychomotor index

experiment were displayed on the screen. This was also handled as an interactive session between the subject and the computer. The instructions appear in Figure 3 as they were presented to the subject. Although the answer given by the subject in Figure 3 shows "yes" to his understanding of the instructions, had he typed "no," he would have been allowed to see the instructions again. The subject was not limited to the number of times he could see the instructions before beginning the learning experiment.

## The Experiment

Each completion of the two sections of the experiment, the acquisition phase and the recall phase, was defined as one trial. The sequential viewing of all 15 items or facts comprised the acquisition phase. Immediately following this phase, test questions, which are shown in Table 2, were presented by the computer to measure the subject's level of learning. This test was the recall phase. The test consisted of one question for each fact and retained the presentation order of the acquisition phase.

After all 15 questions in trial 1 were presented, the items which were answered correctly were removed from the re-presentation of the list in trial 2. Presentation of facts and questions continued by this dropping out procedure until the subject successfully answered all 15 items. This was defined as the criterion for total learning. The total 15-item test was then re-presented to measure the subject's immediate retention of the material.

Unlike the acquisition phase in which the subjects had controlled exposure times, all subjects were self-paced during the recall phase.

Computer: The following is the list of facts for you to commit to memory. You will be presented the facts and then a list of questions about these facts. The facts you miss and the questions pertaining to them will be re-presented until you have answered every question correctly. Is this clear?

Student: Yes

Computer: Finally, after you have given the correct answer to each question you will be given the entire test again. It is <u>very</u> important that you do as well as possible on this one as it is the measure of your retention of these facts. All set?

Student: Yes

Computer: Ok, we're off...

Fig. 3.--Interactive session for instructions to learn

TABLE 2

EXPERIMENTAL TEST AND MINIMUM ACCEPTABLE ANSWERS

| Question Number | Question | Key-Lettered Answer |
|---|---|---|
| 1 | A disease caused by Vitamin A deficiency is night blindness, or what? | nyctalopia |
| 2 | Excess amounts of Vitamin A concentrate might cause what disease? | lipemia |
| 3 | What coenzyme is formed when thiamine combines with phosphates? | cocarboxlas |
| 4 | What is the daily requirement of thiamine? | 0.4 - 1 incl. |
| 5 | Which vitamin is responsible for the respiratory enzyme system? | ribofla |
| 6 | Name one symptom of riboflavin deficiency. | vascul cornea |
| 7 | Diphosphopyridine nucleotide, the coenzyme I of niacine, is also known by what name? | coxymase |
| 8 | What is niacin's coenzyme II, triphospho-pyridine, sometimes called? | coferment |
| 9 | A deficiency of niacin might cause what disease? | pelgra |
| 10 | What is the infant requirement of Vitamin C? | 30 |
| 11 | What is the content of Vitamin C in breast milk? | 4 - 7 incl. |
| 12 | Vitamin D is responsible for the metabolism of what? | phos calc |
| 13 | What beside sunshine and fish liver oils is a good source of Vitamin K? | radiat food |
| 14 | Which vitamin is an essential part of prothrombin formation? | k |
| 15 | What is a source of Vitamin K? | clorphl |

This particular part of the design was unique and not found in traditional verbal learning experimentation concerning the Bugelski hypothesis. The investigator considered this an opportunity to utilize the power of the 1500 system to measure certain behavioral characteristics that the subjects portrayed in responding and to relate them to the four different methods of presentation. This procedure may be criticized for allowing excess time for the subject to rehearse the criterion material during the recall phase. However, the situation was controlled by withholding immediate feedback to the subjects' answers and thereby making rehearsal ineffective. The feedback was delayed until the acquisition phase of the next trial when the unlearned facts were re-presented.

## Other Distinctions

Since this experiment was defined as a study in concept acquisition, it was considered important that the evaluation of the learning during the recall phase not be handicapped by spelling mistakes and/or typing errors. The instructor provided the minimum acceptable letters in sequence considered to constitute the correct answer for each question. The computer was effectively programmed to analyze the subject's response for these certain letters in their particular order and to ignore other intervening letters. In applications of CAI this is known as the "key-letter function." In the verbal learning laboratory this is known as accepting the "functional" response as correct. This has been a topic of research in paired-associate learning (Underwood and Schulz, 1960), and the consensus is that the acceptance of the functional (non-specific) response as correct, in place of the nominal (specific) response, transcends many of the problems of strict stimulus-response associations.

CHAPTER IV

RESULTS AND DISCUSSION

In this chapter, the analysis of the data will be discussed under five sections: (1) Psychomotor index; (2) Number of trials to criterion per group; (3) Group exposure times of the experimental material; (4) Measures of retention; and (5) Group response times for the experimental test.

Psychomotor Index

The psychomotor index was devised on the assumption that it would provide a method of obtaining a measure of the student's ability to function at the terminal. Typing was a psychomotor task, and the individual's skill was one of the obvious behaviors that the psychomotor index purported to measure. Other important elements such as aggressiveness with the equipment and the ability to learn from the computer were also factors considered to be an intrinsic part of this index. The following analysis of the psychomotor index data with respect to the measures of interest in the study (i.e., total exposure time per question and total response time) indicated that it was not an effective criterion against which to measure the kinds of behavior evaluated in this research.

A one-way analysis of variance was used to determine differences which might have existed between the means of the psychomotor indices of the four groups. Table 3 provides the means and the standard deviations of the four groups, and Table 4 shows the analysis of variance for their psychomotor indices. The analysis of variance clearly shows no significant difference between the groups at the .05 level. Duncan's new multiple range test[1] verified that there were no differences between groups.

TABLE 3

MEAN PSYCHOMOTOR INDEX FOR EACH GROUP
WITH STANDARD DEVIATION AND RANGE

| Time Group | N | Mean | Standard Deviation | Range Minimum | Range Maximum |
|---|---|---|---|---|---|
| Minimum | 12 | 57.42 | 14.25 | 27.0 | 79.0 |
| Comfortable | 13 | 67.39 | 29.64 | 36.0 | 150.0 |
| Excessive | 15 | 62.07 | 37.72 | 24.0 | 173.0 |
| Self-Pacing | 12 | 58.58 | 22.34 | 32.0 | 110.0 |

TABLE 4

ONE-WAY ANALYSIS OF VARIANCE FOR PSYCHOMOTOR INDEX

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Total | 51 | 38946.98 | | |
| Groups | 3 | 757.14 | 252.38 | .317 |
| Error | 48 | 38189.84 | 795.62 | |

[1]For each analysis of variance, a multiple comparisons test was required to determine significant differences between groups. Duncan's new multiple range test was chosen as the appropriate test to provide this information. A FORTRAN program following Steel and Torrie's (1960) procedure was used to perform this test for the results of a completely randomized design analysis of variance such as in Table 4.

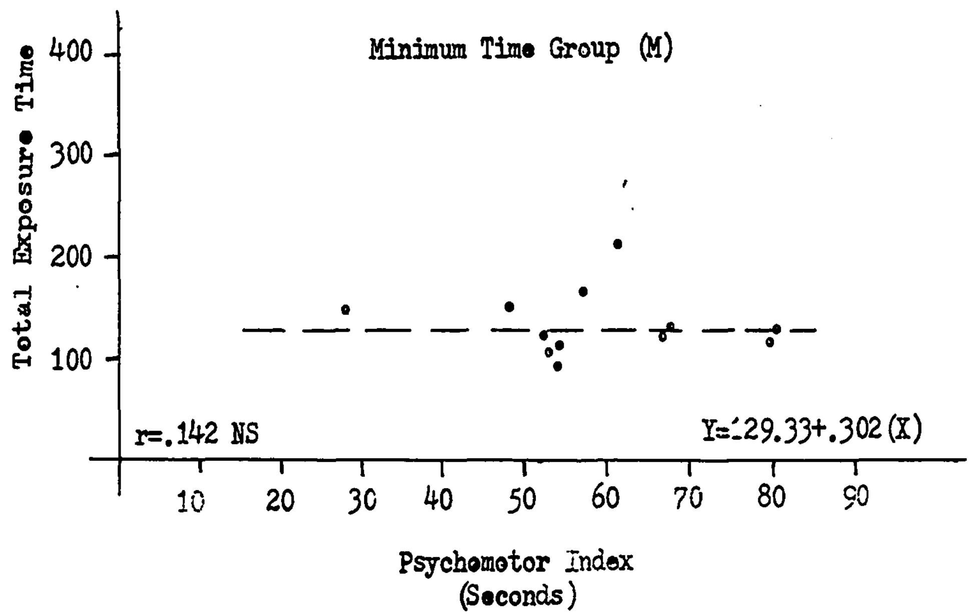
Correlation tests were performed in an attempt to establish a relationship between the psychomotor index and the total exposure time. The results are listed in Table 5. As shown, none of the correlation coefficients was significant at the .05 level. The relationship could not be established between the two measures.

TABLE 5

CORRELATION COEFFICIENTS BETWEEN PSYCHOMOTOR INDEX AND TOTAL EXPOSURE TIME

| Time Group | N | r | Significance (.05) |
|---|---|---|---|
| Minimum | 12 | .14 | NS |
| Comfortable | 13 | .20 | NS |
| Excessive | 15 | .12 | NS |
| Self-Paced | 12 | .13 | NS |

Figures 4 and 5 were plotted to provide visual evidence of the behavior patterns portrayed by the groups that concerned the psychomotor index (X-axis) and the total exposure time (Y-axis). Each of the graphs represents one group in the study. It is particularly interesting to notice the excessive time group (E) and the self-pacing group (SP). The apparent trend of the points on these two graphs seems to be more pronounced than on the graphs for the minimum time group (M) and the comfortable time group (C). Since the former two groups were operating under conditions of more time and less stress than the latter two groups, one might surmise that the spread indicates some validity in the psychomotor index. Groups M and C's lack of functional relationship between the two variables might indicate that other motivational forces were compromising



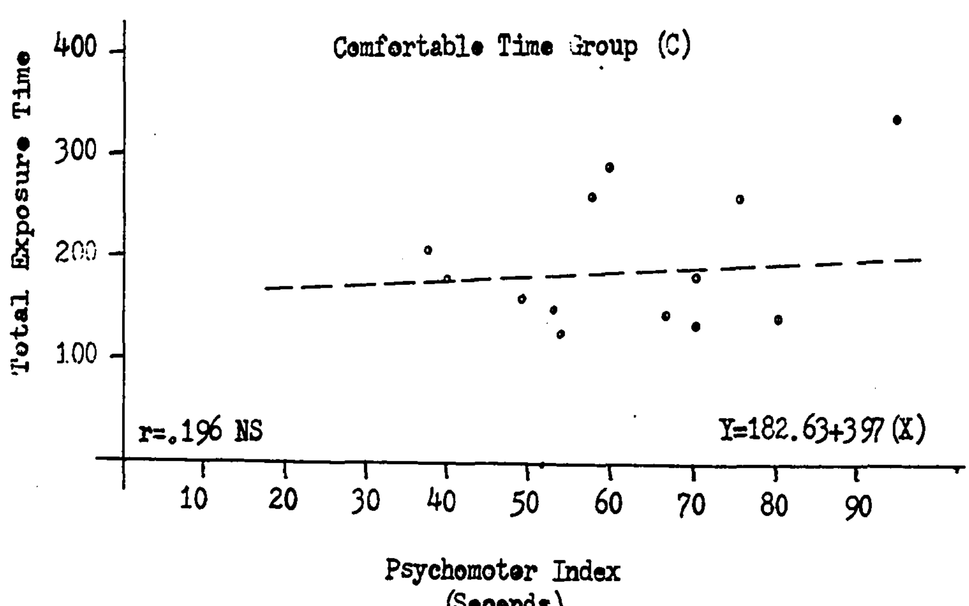
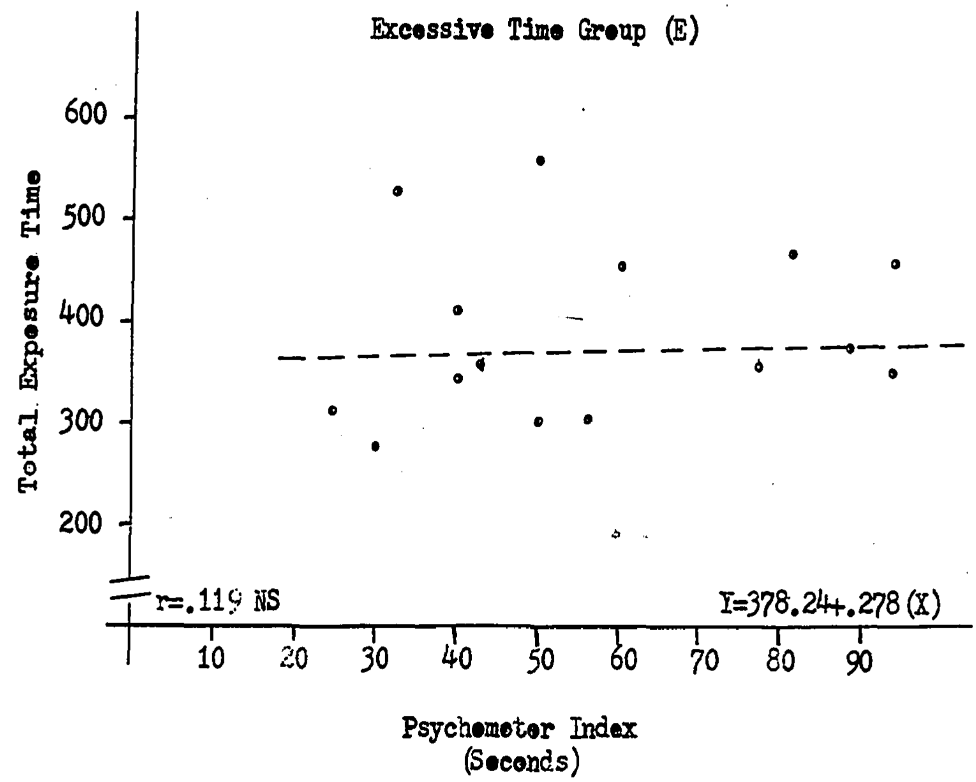




Fig. 4.--Linear regression lines for total exposure time versus psychomotor index - Groups M and C



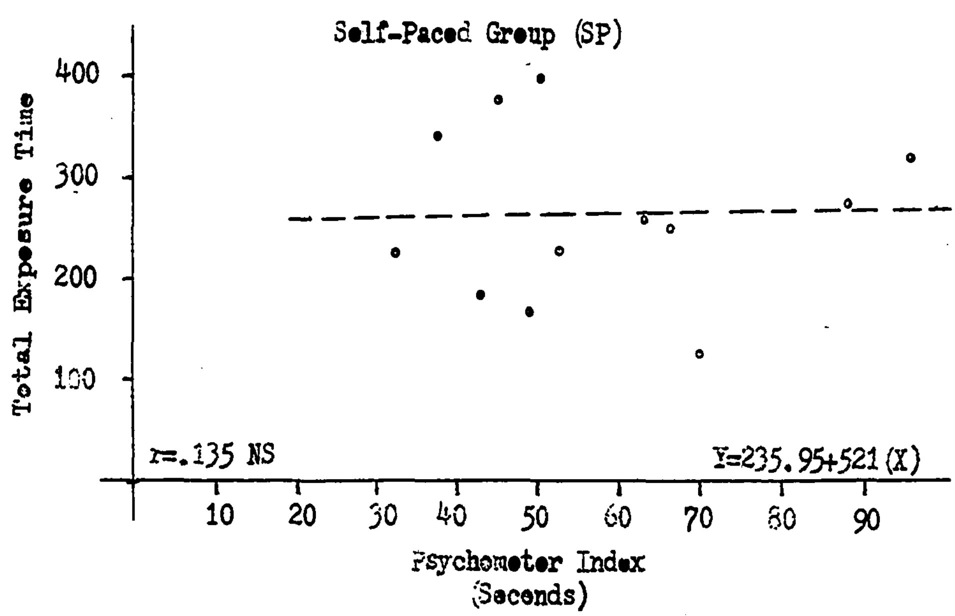




Fig. 5.--Linear regression lines for total exposure time versus psychometer index - Groups E and SP

whatever effect the psychomotor index had in controlling the subjects' behavior. However, the non-significance of the correlation coefficients did not support these conclusions.

Regression coefficients were performed for each of the four groups, and the regression line was plotted on each graph with its appropriate equation. Regression coefficients for M and C were not significant at the .05 level, but the regression coefficients for E and SP were statistically significant. This supports the observation that there was a trend which existed in the relationship between total exposure time and the psychomotor index measures for the groups which were allowed as much time as they liked in functioning at the terminal.

While the correlation between psychomotor index and total exposure time was not statistically significant, the total response time and psychomotor index were apparently related for some groups. Consideration of this relationship was thought to be especially meaningful in the total response time of the subject on trial 1 when the criterion material was still unfamiliar. Conceivably, an early relationship could provide more information concerning the validity of the psychomotor index as a measure of functional ability. If, however, the subjects' behavior during the performance of the task that comprised the psychomotor index was not correlated to the more intensive task of answering the list of 15 questions on the first trial, it would be safe to eliminate the psychomotor index from further consideration as an important measure of functional ability. Table 6 presents correlation coefficients for the psychomotor index and the total response times on trial 1.

The significance of C and E and the absence of significance of M and SP might suggest that different motivational forces which interfered

TABLE 6

CORRELATION COEFFICIENTS BETWEEN PSYCHOMOTOR INDEX
AND TOTAL RESPONSE TIMES - TRIAL 1

| Time Group | N | r | Significance (.05) |
|---|---|---|---|
| Minimum | 12 | .13 | NS |
| Comfortable | 13 | .73 | * |
| Excessive | 15 | .82 | * |
| Self-Paced | 12 | .07 | NS |

with the effects measured by the psychomotor index were acting on the latter two groups. One might speculate that the subjects in M were operating under stress during trial 1 because they knew very few answers. The design of the experiment was such that, at the beginning of the recall phase of trial 1, M had seen all of the material only long enough to read it. Table 7 shows that the mean number of items recalled by M on trial 1 was 4.167. Table 8 is an analysis of variance showing significant differences between groups concerning mean items recalled on trial 1 ($F = 8.298$). Table 9 shows where these differences occurred. It is curious that, compared to the other groups, M recalled significantly fewer items after the first exposure to the experimental material and still took the least amount of time (333.71 seconds) to go through the first recall phase. It is possible that the subjects in M were so unfamiliar with the material after only one fast exposure that they did not anticipate the answers to many of the questions.

On the other hand, it may be surmised that SP was more comfortable with the equipment by this time because of having interacted with the system (by pressing the space bar) during the acquisition phase. Tables 7, 8, and 9 support the idea that SP learned more during trial 1 than the other groups. The average number of the 15 items recalled by SP on the first trial was 9.083. Apparently, during trial 1 they were not experiencing anxiety over the machinery nor any appreciable confusion or frustration due to unfamiliarity with the criterion material.

TABLE 7

MEAN NUMBER OF ITEMS RECALLED AND MEAN RESPONSE TIMES FOR EACH GROUP - TRIAL 1

| Time Group | N | Mean Items Recalled | Mean Response Times (Seconds) |
|---|---|---|---|
| Minimum | 12 | 4.167 | 333.71 |
| Comfortable | 13 | 6.846 | 513.04 |
| Excessive | 15 | 7.000 | 379.69 |
| Self-Paced | 12 | 9.083 | 388.72 |

TABLE 8

ONE-WAY ANALYSIS OF VARIANCE FOR ITEMS RECALLED - TRIAL 1

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Total | 51 | 428.67 | | |
| Group | 3 | 146.40 | 48.80 | 8.30* |
| Error | 48 | 282.28 | 5.88 | |

*$p < .005$

TABLE 9

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
NUMBER OF ITEMS RECALLED - TRIAL 1

| Group | Significantly Different From |
|---|---|
| Minimum | C, E, SP |
| Comfortable | M, SP |
| Excessive | M, SP |
| Self-Paced | M, C, E |

At the time that the immediate retention test was given, total learning of the experimental material had occurred for all groups. It may be assumed that by this time all subjects had become accustomed to the material presentation via the CRT. Table 10 shows the correlation coefficients between the psychomotor index and total response time on the immediate retention trial. Only group E had retained a significant relationship to this point. In view of their excessive exposure times, it is possible that their responding patterns were not influenced by the system. Groups M and C were perceptually stimulated by relatively rapid changing of items on the screen, and SP was having to press the space bar to change the items. E, however, was a passive group with little opportunity for interaction.

The psychomotor index or some comparable index is a potentially valuable measure to be considered in future research in computer-assisted instruction. Refinements of the technique used here or, perhaps, complete revision of it would be necessary. This particular index was too inconsistent between groups and over different situations to be of any value.

The results of the analysis concerning the psychomotor index contraindicated its use as a covariate for the remaining portion of the data analysis.

TABLE 10

CORRELATION COEFFICIENTS BETWEEN PSYCHOMOTOR INDEX AND TOTAL RESPONSE TIMES - IMMEDIATE RETENTION TEST

| Time Group | N | r | Significance (.05) |
|---|---|---|---|
| Minimum | 12 | .27 | NS |
| Comfortable | 13 | .20 | NS |
| Excessive | 15 | .88 | * |
| Self-Paced | 12 | .42 | NS |

<u>Trials to Criterion</u>

To provide support for the total time hypothesis, the number of trials that a subject required to reach criterion was related to his treatment group. Since M was exposed to the experimental material only one-half as long per item as C, it was assumed that they would have required approximately twice as many exposures as C to reach criterion. Indeed, for the situation to have been entirely symmetrical, the mean trials to criterion for M should have been approximately twice as great as the trials to criterion for C and four times as great as the trials to criterion for E. Table 11 shows the mean trials to criterion, and this symmetry obviously did not materialize.

An analysis of variance for the number of trials to criterion was performed to determine the differences between groups, questions, and if the two interacted. The design was a factorial design with repeated measures and followed the procedure outlined in a standard statistical text (Winer, 1962, pp. 374-378).[1] The procedure allowed each subject to be treated as observed under more than one treatment condition (question). Clearly, the F values as shown in Table 12 are significant. There were obviously group differences, question differences, and the two were interacting. This confirmed the implied hypothesis that there was a difference between the groups concerning the number of trials required to reach criterion. This finding was compatible with the research findings of the total time hypothesis in the verbal learning laboratory, but the lack of symmetry discussed earlier provided a good indicator that the end of the compatibility was in sight. The Duncan's new multiple range test was performed to isolate the differences. The results appear in Table 13.

TAPLE 11

MEAN NUMBER OF TRIALS TO REACH CRITERION PER GROUP WITH RANGES*

| Time Group | N | Mean | Range | |
|---|---|---|---|---|
| | | | Minimum | Maximum |
| Minimum | 12 | 7.25 | 3.0 | 13.0 |
| Comfortable | 13 | 4.92 | 2.0 | 9.0 |
| Excessive | 15 | 4.33 | 2.0 | 7.0 |
| Self-Paced | 12 | 3.08 | 2.0 | 7.0 |

*The error mean squares (variance) for computing the pooled standard error for these mean values is shown in Table 12 as subject within groups mean squares.

[1]In several cases throughout this paper an analysis of variance was used for discrete data. In each of these cases a frequency distribution was compiled to establish normalcy of the data in each group. Since normalcy appeared to exist, the analysis of variance was chosen for its interpretational and computational ease.

TABLE 12

ANALYSIS OF VARIANCE FOR TRIALS TO CRITERION

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Between Subjects | | | | |
| Groups (A) | 3 | 176.98 | 58.96 | 14.14* |
| Subjects within Groups | 48 | 200.03 | 4.16 | |
| Within Subjects | | | | |
| Questions (B) | 14 | 486.49 | 34.74 | 25.44* |
| AB | 42 | 188.28 | 4.48 | 4.28* |
| Question x Subjects within Groups | 672 | 917.65 | 1.36 | |

*$p < .005$

TABLE 13

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
MEAN NUMBER OF TRIALS TO CRITERION

| Group | Significantly Different From |
|---|---|
| Minimum | C, E, SP |
| Comfortable | M, SP |
| Excessive | M, SP |
| Self-Paced | M, C, E |

Indications are that C and E were behaving the same concerning the trials it took to reach criterion. The two were not shown to be significantly different by the multiple range test and the means are very close to equality (Table 11). This is consistent with the previous discussions concerning mean items recalled on trial 1 and also the correlation of total response time on trial 1 with the psychomotor index. SP was the most effective and M the least effective concerning the number of trials taken to reach criterion. SP and E were significantly different in this respect.

An interesting point concerns the source of variation in the group-question interaction. The high significance of this interaction suggests that the groups were not treating the questions in the same way. This clue to behavioral differences might have suggested differences between free cognitive styles and structured lessons. Figure 6 was developed to provide insight into this situation and shows the interaction that occurred between questions and groups with respect to the number of trials to criterion. Table 14 is provided as a reference for the analysis of the question differences by groups. Primarily, the graph shows that the interaction indicated by the significant F value in the AOV was the result of the accumulation of many small differences and was not indicative of significantly different behavior. The graph conveys the idea that the four groups were behaving similarly concerning the trials to criterion.

It may be noticed that M's behavior was the most extreme concerning trials to criterion. This was expected since they were allowed so little time per item to acquire the information. Even this extreme behavior is consistent with the other groups on a pattern basis. One would surmise from the variation in the line created by M's behavior that questions 3, 7, and 13 were questions with particular content difficulty. Question 1 was probably giving M difficulty due to the fact that the built-in time that the computer allowed this to be shown on the screen approached the lower perceptual threshold of many subjects. Some attention

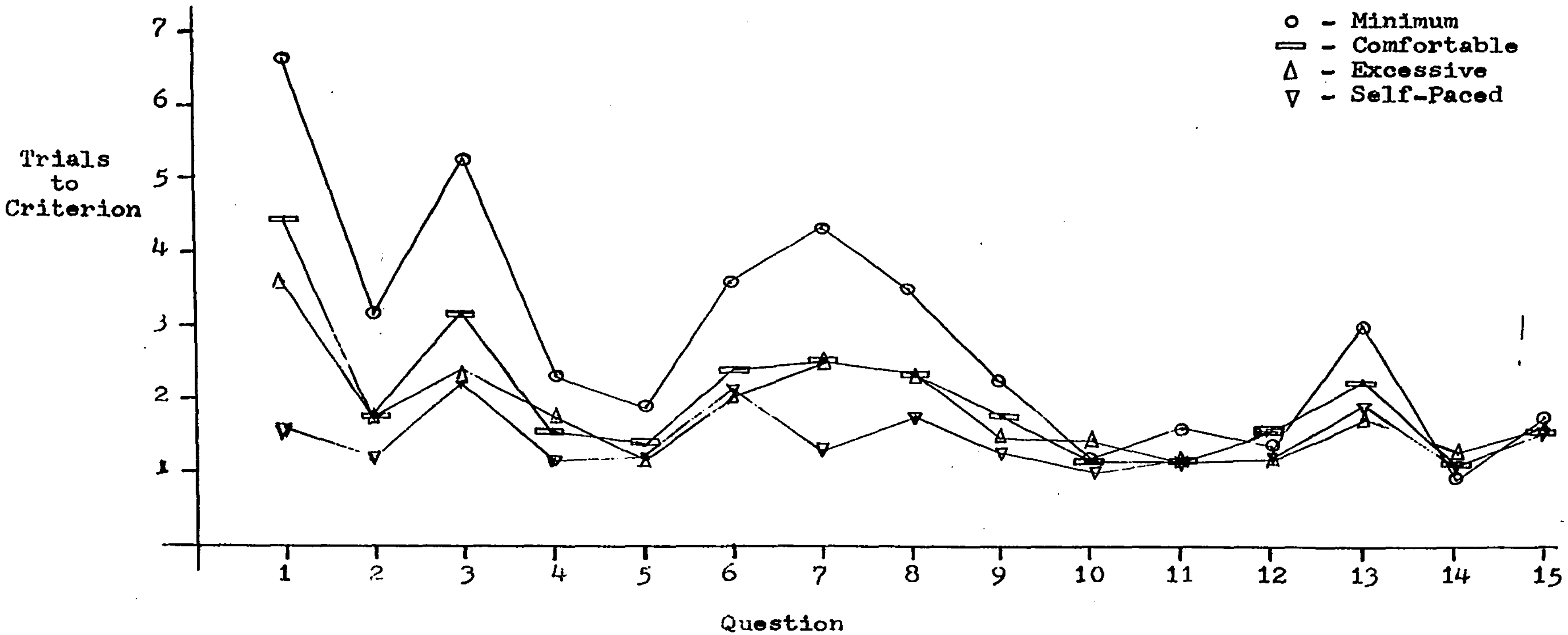



Fig. 6.--Plot of intergroup interaction: trials to criterion versus question

TABLE 14

MEAN NUMBER OF TRIALS TO CRITERION PER ITEM

| Item Number | Group | | | |
|---|---|---|---|---|
| | Minimum | Comfortable | Excessive | Self-Paced |
| 1 | 6.67 | 4.38 | 3.60 | 1.67 |
| 2 | 3.08 | 1.85 | 1.87 | 1.17 |
| 3 | 5.17 | 3.08 | 2.47 | 2.25 |
| 4 | 2.25 | 1.69 | 1.87 | 1.17 |
| 5 | 1.92 | 1.54 | 1.27 | 1.33 |
| 6 | 3.58 | 2.31 | 2.00 | 2.08 |
| 7 | 4.25 | 2.61 | 2.67 | 1.17 |
| 8 | 3.50 | 2.31 | 2.27 | 1.83 |
| 9 | 2.16 | 1.85 | 1.40 | 1.25 |
| 10 | 1.25 | 1.31 | 1.53 | 1.08 |
| 11 | 1.75 | 1.31 | 1.33 | 1.25 |
| 12 | 1.58 | 1.69 | 1.53 | 1.50 |
| 13 | 3.00 | 2.38 | 1.80 | 1.90 |
| 14 | 1.00 | 1.08 | 1.33 | 1.08 |
| 15 | 1.92 | 1.69 | 1.80 | 1.83 |

should also be given to the idea that the amount of retroactive inhibition[1] was possibly greater for this item than for any other in the list. It may be argued that the subject was required to participate in more

[1]Retroactive inhibition may be defined as the learning of material that tends to negatively affect the subject's recall of similar material previously learned.

uninterrupted study time over the criterion material between the time this item was presented and the time the question was asked about it. Perhaps the most plausible explanation lies in the fact that the answer for question 1 was "nyctalopia" and the nominal response was required (see Table 2). This presented the possibility of added trials for subjects misspelling the response. This seemed more likely for group M than for any other group due to their short exposure time to the word per trial.

The other three groups plotted on the chart supported the notion that, given the proper amount of time, the behavior of the subjects will be very much the same. They show no extremes, as M does, and the patterns formed by the lines are similar. The self-paced control group's behavior was particularly supportive of this idea. At virtually every point on the chart, SP had a lower trials-to-criterion score than the other three groups. This may be considered in terms of SP's having been allowed to determine the necessary study time for each item and decide when learning had occurred.

## Group Exposure Times

The Bugelski total time hypothesis primarily concerns the amount of time it takes for a subject to acquire a specified amount of material. Assuming adequate experimental controls, the validity of this hypothesis at the concept level can be accepted or rejected on the basis of the mean exposure times recorded across the various treatment conditions. Table 15 shows the mean amount of time spent by each group on each question, and the values of the mean total exposure times per group are shown in Table 16.

TABLE 15

MEAN EXPOSURE TIME PER ITEM (IN SECONDS)

| Item Number | Minimum | Group<br>Comfortable | Excessive | Self-Paced |
|---|---|---|---|---|
| 1 | 10.00 | 13.15 | 21.60 | 38.68 |
| 2 | 9.25 | 11.08 | 22.40 | 14.42 |
| 3 | 15.50 | 18.46 | 29.60 | 33.42 |
| 4 | 5.85 | 8.80 | 19.41 | 13.47 |
| 5 | 6.13 | 9.85 | 16.21 | 12.97 |
| 6 | 8.96 | 11.54 | 20.00 | 20.87 |
| 7 | 25.35 | 32.43 | 66.20 | 42.17 |
| 8 | 17.15 | 22.62 | 45.09 | 20.63 |
| 9 | 6.72 | 11.45 | 17.36 | 6.75 |
| 10 | 4.25 | 8.83 | 20.85 | 5.60 |
| 11 | 7.35 | 10.97 | 22.40 | 15.17 |
| 12 | 8.52 | 12.86 | 23.31 | 10.52 |
| 13 | 13.80 | 21.93 | 33.12 | 15.32 |
| 14 | 3.60 | 7.75 | 16.32 | 6.03 |
| 15 | 5.75 | 10.15 | 21.60 | 9.44 |

Table 17 presents an analysis of variance which was performed on data presented in Table 15 utilizing the same factorial design found on page 36. Clearly, there is significance concerning group differences, question differences, and group-question interaction. Duncan's new multiple range test was performed on the mean exposure times and indicated significant differences between each of the groups (Table 18). The total time hypothesis was, therefore, rejected as a valid hypothesis at the concept level.

TABLE 16

MEAN TOTAL EXPOSURE TIME (IN SECONDS) PER GROUP WITH RANGES

| Time Group | N | Mean | Range Minimum | Range Maximum |
|---|---|---|---|---|
| Minimum | 12 | 146.67 | 99.3 | 213.3 |
| Comfortable | 13 | 209.37 | 150.6 | 340.6 |
| Excessive | 15 | 395.48 | 283.2 | 566.4 |
| Self-Paced | 12 | 266.47 | 137.7 | 401.4 |

TABLE 17

ANALYSIS OF VARIANCE FOR EXPOSURE TIME BY QUESTION FOR EACH GROUP

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Between Subjects | | | | |
| Groups (A) | 3 | 28297.23 | 9426.41 | 26.45* |
| Subjects within Groups | 48 | 17105.00 | 356.35 | |
| Within Subjects | | | | |
| Questions (B) | 14 | 55046.06 | 3931.86 | 32.38* |
| AB | 42 | 16331.70 | 388.85 | 3.20* |
| Question x Subjects within Groups | 672 | 81595.02 | 121.42 | |

*p < .005

TABLE 18

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
MEAN EXPOSURE TIME PER GROUP

| Group | Significantly Different From |
|---|---|
| Minimum | C, E, SP |
| Comfortable | M, E, SP |
| Excessive | M, C, SP |
| Self-Paced | M, C, E |

The differences noted might be explained by a concept developed in the verbal learning laboratory called stage analysis (Underwood, Runquist, and Schulz, 1959). Stage analysis simply says that the acquisition of criterion material is accomplished in two stages. The first stage is the learning of the response item in the list, and the second stage is the association or "hookup" of the response item to the stimulus item. Stage analysis purports that the subject becomes involved in this two-stage process every time the material is perceived. If this is true at the concept level, the subjects (M) in the present experiment stimulated by repetitive exposures should have learned faster as, indeed, they did.

The stage-analysis idea is not to be confused with the difference between nominal and effective time which is one of the conditions of the total time hypothesis. Effective time is defined as the time the subject attends to the task. There is no evidence that the subjects in C and E were not using their time as effective time. The repetitive exposures which facilitated the learning of M concerned method (rate) of stimulating the rehearsal process. For M this method involved one rehearsal per item with repetitive exposures of the item. On the other hand, C and E were providing their own stimulus for repetitive rehearsal during the time allowed each item. The data did seem to say that M's method was the most conducive to rapid acquisition of the material.

An analysis of the group-question interaction plots revealed behavioral characteristics of the groups during the acquisition of the individual items. Figure 7 is a picture of the behavior of M, C, and E and clearly shows the similarity of the patterns of their exposure times. While the similar behavior of these three groups over the whole sequence of questions was remarkable, it is equally as dramatic to observe on Figure 8 how differently the self-paced group behaved concerning the acquisition time per question. The graphs clearly show that the significance of the interaction source was due to the behavior of the subjects in SP. Specifically, Figure 8 shows SP as the group with the longest exposure time for items 1 and 3. It might be recalled that Figure 6 shows SP as the fastest group concerning trials to criterion for these particular items.

Several explanations were considered for this paradox. One was that SP used the first item in the list to become adjusted to their environment. This was a condition not afforded the other groups since they had no control over the exposure time. This explanation, while logical, was confounded by the fact that item 3 as well as 1 was involved in the same paradoxical situation while item 2 was not. The retroactive inhibition phenomenon mentioned earlier was also considered as a possible explanation because of the long period of uninterrupted study time between the early items in the list and the test. This explanation assumed that

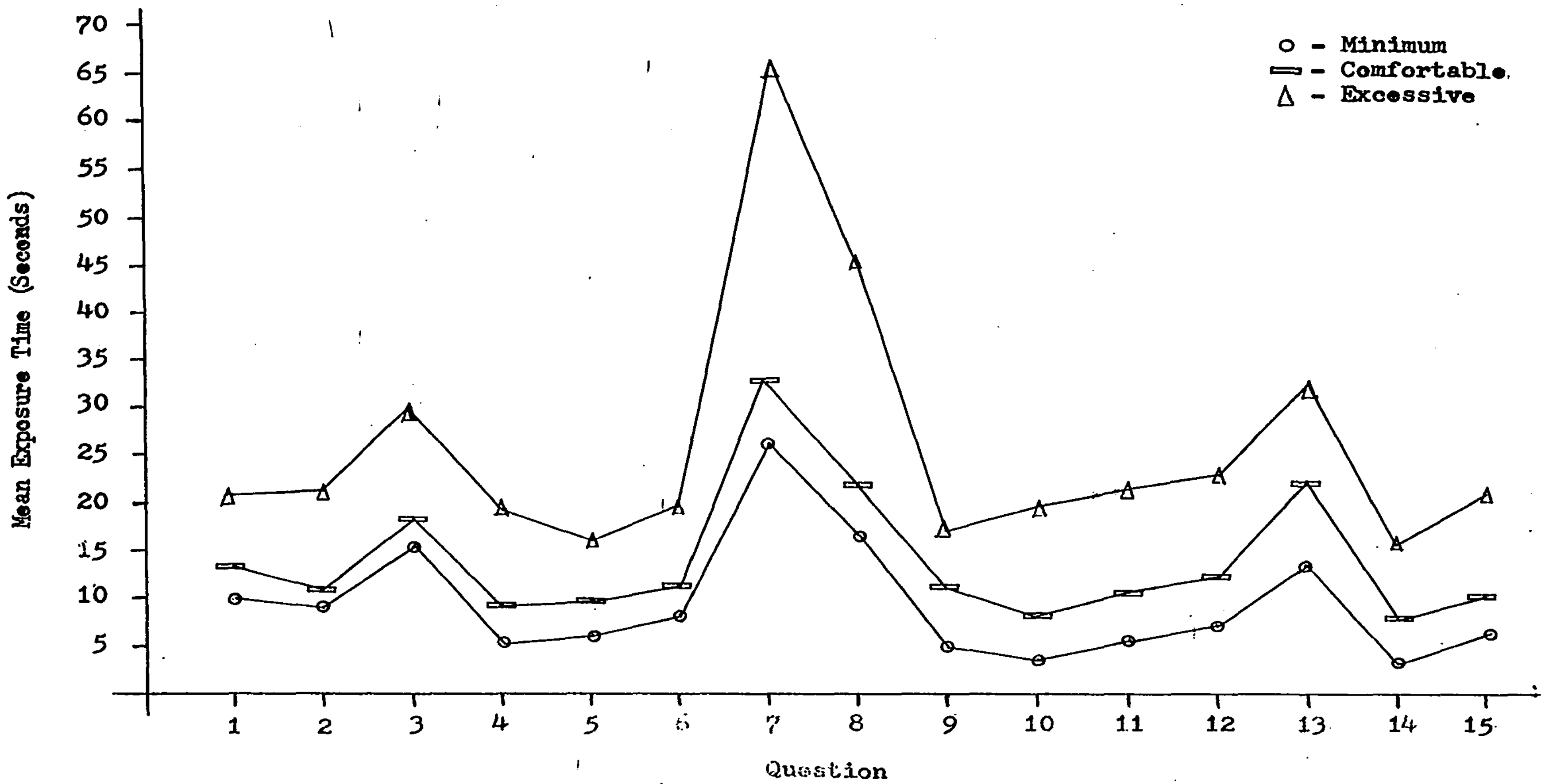



Fig. 7.--Plot of interaction of groups M, C, and E: Mean exposure time versus question

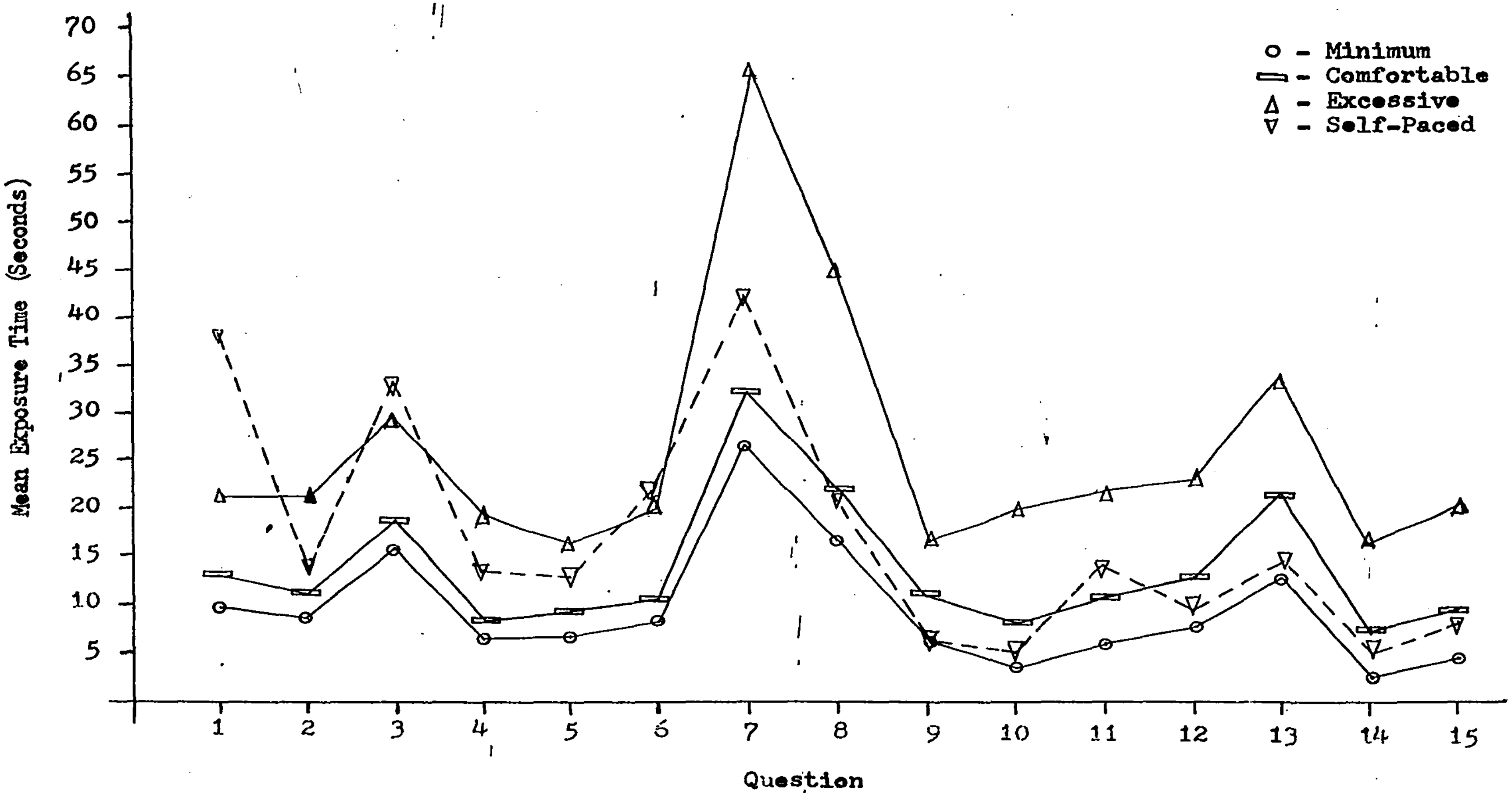



Fig. 8.--Plot of interaction of all groups: Mean exposure time versus question

item 2 was easier than items 1 and 3, an assumption confirmed by the instructor of the subject matter. However, the operation of retroactive inhibition would have caused the subject to respond incorrectly and, consequently, require more trials to reach criterion.

The fact that SP acted as the self-paced control group in the design of the experiment possibly accounted for their differential behavior on items 1 and 3. The implication is that the subject's time was used most efficiently when he was allowed to decide which items were most difficult and required more study time. It may be recalled from the discussion of method that the times assigned to the criterion material for M, C, and E were based on reading speed alone. These three groups were seeing the study material for time periods based solely on the number of words in the statement, and no attention was paid to the complexity of the concept. The subjects in the self-pacing group were not faced with this restriction. They read the material, observed the complexities, and made decisions concerning which concepts required more rehearsal time. This is the most plausible explanation of SP's longer exposure time and fewer trials to criterion for items 1 and 3.

## Measures of Retention

It may be recalled from Chapter 3, the discussion of method, that upon reaching criterion the subjects were re-presented the complete test. Their performance on this re-examination provided a measure of immediate retention. Tables 19, 20, and 21 provide means, F values, and the Duncan's new multiple range test results.

The F value (1.9) concerning overall difference is not statistically significant. This indicated the acceptance of the hypothesis

TABLE 19

MEAN NUMBER OF ITEMS RECALLED - IMMEDIATE RETENTION TEST

| Time Group | N | Mean | Range Minimum | Range Maximum |
|---|---|---|---|---|
| Minimum | 12 | 12.00 | 8.0 | 15.0 |
| Comfortable | 13 | 12.92 | 7.0 | 15.0 |
| Excessive | 15 | 13.67 | 11.0 | 15.0 |
| Self-Paced | 12 | 13.25 | 9.0 | 15.0 |

TABLE 20

ONE-WAY ANALYSIS OF VARIANCE FOR ITEMS RECALLED - IMMEDIATE RETENTION TEST

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Total | 51 | 182.00 | | |
| Treatment | 3 | 19.49 | 6.50 | 1.9 |
| Error | 48 | 162.51 | 3.39 | |

TABLE 21

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS: NUMBER OF ITEMS RECALLED - IMMEDIATE RETENTION TEST

| Group | Significantly Different From |
|---|---|
| Minimum | E |
| Comfortable | --- |
| Excessive | M |
| Self-Paced | --- |

which stated that there would be no difference between the groups' retention of the material. However, the multiple range test results showed a significant difference between the immediate retention of M and E. In percentage figures, M retained 80 per cent of the items learned and E retained 91 per cent (Table 22).

TABLE 22

PERCENTAGE OF EXPERIMENTAL ITEMS RETAINED
ON IMMEDIATE RETENTION TEST

| Group | Percent |
|---|---|
| Minimum | 80 |
| Comfortable | 86 |
| Excessive | 91 |
| Self-Paced | 88 |

The difference in retention was accredited to the lack of opportunity afforded M to overlearn any of the material. M's perceptual threshold was approached with the presentation of each item, a condition that allowed no extra rehearsal time. On the other hand, E was in a situation that encouraged overlearning. The fact that E apparently did overlearn indicates that the excessive exposure rate was being used as effective time. The hypothesis can, therefore, be accepted with the exception of the difference that exists between M and E.

## Response Time Analysis

The rationale for allowing all subjects to pace themselves through the recall phase was to gather information relating their behavior to different presentation conditions. As they answered questions during this phase, their response times were recorded by the computer. These recorded response times provided information on: (1) the conditioning of the subjects' responding times by the pace of the acquisition phase; (2) the total time on the test; and (3) the groups' varied abilities to clearly separate the acquisition phase from the recall phase in the learning situation.

The mean total response times for the four groups are found in Table 23. An analysis of variance was performed and disclosed a significant difference (F = 4.25) between groups concerning response times per question. This analysis is shown in Table 24. Duncan's new multiple range test was performed and is summarized in Table 25.

TABLE 23

MEAN TOTAL RESPONSE TIME PER GROUP WITH RANGES

| Time Group | N | Mean | Range | |
|---|---|---|---|---|
| | | | Minimum | Maximum |
| Minimum | 12 | 1001.2 | 198.1 | 506.7 |
| Comfortable | 13 | 1250.4 | 318.9 | 1010.7 |
| Excessive | 15 | 939.8 | 188.9 | 867.9 |
| Self-Paced | 12 | 751.3 | 215.1 | 488.5 |

TABLE 24

ANALYSIS OF VARIANCE FOR TOTAL RESPONSE TIME PER QUESTION

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Between Subjects | | | | |
| Groups (A) | 3 | 97859.40 | 32619.79 | 4.25* |
| Subjects within Groups | 48 | 371774.61 | 7745.30 | |
| Within Subjects | | | | |
| Questions (B) | 14 | 544042.07 | 38860.14 | 7.43* |
| AB | 42 | 105416.86 | 2509.92 | 1.77 |
| Question x Subjects within Groups | 672 | 952016.12 | 1416.69 | |

*p <.005

TABLE 25

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
TOTAL RESPONSE TIME

| Group | Significantly Different From |
|---|---|
| Minimum | C, SP |
| Comfortable | M, E, SP |
| Excessive | C |
| Self-Paced | M, C |

It was hypothesized that the rate of experimental material presented by the computer would not affect the rate of responding by the subjects. These data did not support this hypothesis. It has been shown that group C had an average of 4.92 trials to criterion and required an average of 1258.38 seconds of response time. M, on the other hand, averaged 7.25 trials in reaching criterion with an average total response time of 1001.7 seconds. The fact that M took 2.33 trials more than C and still required 256.7 seconds less in response time presented a paradox that suggested the existence of a pacing phenomenon which could have implications in future CAI material development.

Before this phenomenon could be accepted, it had to be observed under conditions where response times were not confounded by different levels of learning. The psychomotor index was such a measure where the subjects' level of learning was not a factor. It may be recalled from Table 4 that there was no significant difference between M and C for this response time measure. The immediate retention measure was taken under conditions of equal learning because the groups had both reached the defined criterion. The mean response times for the immediate retention trial are provided in Table 26. An analysis of variance showed no significance (F = 1.79), but Duncan's new multiple range test (Table 28) showed that M's average response time was significantly less than C's on the immediate retention trial. The two groups' apparent adjustments to their individual exposure pace is strong evidence for the existence of the phenomenon.

TABLE 26

MEAN RESPONSE TIMES (SEC.) - IMMEDIATE RETENTION TRIAL

| Time Group | N | Mean | Range Minimum | Range Maximum |
|---|---|---|---|---|
| Minimum | 12 | 221.22 | 142.20 | 310.20 |
| Comfortable | 13 | 313.15 | 186.30 | 582.50 |
| Excessive | 15 | 261.97 | 135.80 | 588.90 |
| Self-Paced | 12 | 257.81 | 138.20 | 446.80 |

TABLE 27

ANALYSIS OF VARIANCE FOR RESPONSE TIME - IMMEDIATE RETENTION TRIAL

| Source | df | SS | MS | F |
|---|---|---|---|---|
| <u>Between Subjects</u> | | | | |
| Groups (A) | 3 | 3684.05 | 1228.01 | 1.79 |
| Subjects within Groups | 48 | 32759.75 | 682.49 | |
| <u>Within Subjects</u> | | | | |
| Questions (B) | 14 | 17947.70 | 1281.97 | 10.76* |
| AB | 42 | 6084.58 | 144.87 | 1.21 |
| Question x subject within Groups | 672 | 80038.96 | 119.10 | |

*p $<.005$

TABLE 28

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
TOTAL RESPONSE TIMES - IMMEDIATE
RETENTION TRIAL

| Group | Significantly Different From |
|---|---|
| Minimum | C |
| Comfortable | M |
| Excessive | --- |
| Self-Paced | --- |

Some consideration was given to the pacing of E since they were also under the control of the computer. It was previously suggested that the subjects' responding behavior apparently was not influenced by such a slow pace. Support for this was obtained by comparing E to the control group of self-pacers (SP). Table 29 summarizes the group differences concerning situations where the subject was required to respond. Since there was no significant difference between E and SP over any of these conditions, it can be surmised that E was behaving like the unpaced control group (SP) during all response situations.

TABLE 29

SUMMARY OF DIFFERENCES AMONG GROUPS IN VARIOUS RESPONDING SITUATIONS

| Group | Significantly Different From | | | |
|---|---|---|---|---|
| | Psychomotor Index | Trial 1 | Immediate Retention Trial | Total Response Time |
| Minimum | --- | C | C, SP | C |
| Comfortable | --- | M | M, E, SP | M, E, SP |
| Excessive | --- | --- | C | C |
| Self-Paced | --- | --- | M, C | C |

Data analysis for the time spent by each group on the total experimental task appears in Tables 30, 31, and 32. Based on this evidence, the hypothesis of no difference between groups concerning total amount of time involved in the experimental task must be rejected.

However, a notable finding was the lack of significant difference between M and SP with respect to the total amount of time to complete the experimental task (Table 32). Although the two groups' total time was the same (Table 30), M performed better than SP on total exposure time (Table 16) and poorer than SP on total response time (Table 23). This suggested that these groups were learning by two different methods: the anticipation method and the recall method.

TABLE 30

MEAN TOTAL TIME ON EXPERIMENTAL TASK (IN SECONDS) PER GROUP WITH RANGES

| Time Group | N | Mean | Range | |
|---|---|---|---|---|
| | | | Minimum | Maximum |
| Minimum | 12 | 1147.33 | 733.0 | 1551.0 |
| Comfortable | 13 | 1467.31 | 855.0 | 2104.0 |
| Excessive | 15 | 1335.00 | 667.0 | 2433.0 |
| Self-Paced | 12 | 1017.77 | 586.0 | 1393.0 |

TABLE 31

ANALYSIS OF VARIANCE FOR TOTAL TIME ON EXPERIMENTAL TASK

| Source | df | SS | MS | F |
|---|---|---|---|---|
| Total | 51 | 8336821.1 | | |
| Group | 3 | 1725103.0 | 575034.3 | 4.17* |
| Error | 48 | 6611718.1 | 137744.1 | |

*p < .005

TABLE 32

DUNCAN'S NEW MULTIPLE RANGE TEST RESULTS
TOTAL TIME ON EXPERIMENTAL TASK

| Group | Significantly Different From |
|---|---|
| Minimum | C |
| Comfortable | M, SP |
| Excessive | SP |
| Self-Paced | C, E |

The difference between these two methods is discussed by Kausler (1966, pp. 136-137). Basically, the anticipation method requires that the subject learn at virtually the same time he is called on to perform. He is being asked to anticipate answers which he does not know. This dual responsibility of learning and performing should generate considerable interference in the absence of feedback and reduce the subject's overall performance. The recall method provides a clear separation of learning (study) and performance (test) which transcends the source of confounding present in the anticipation method.

The subjects in SP, who were in complete control of the experimental situation, were able to clearly separate the acquisition phase from the recall phase. Their learning took place as the facts were presented. Then, as the subjects saw the questions, they recalled those facts already learned. M subjects, on the other hand, were not allowed time to complete their learning during the acquisition phase. Consequently, when they came to the testing phase, they were forced into the position of anticipating rather than recalling the answers. One might expect the lack of immediate feedback in the testing phase to have caused more debilitation of learning for M than SP. These data did not support this. While some research in verbal learning has indicated no deleterious effects from delayed information feedback (Bilodeau and Bilodeau, 1958), it is suspected that group M might have been even more superior in total time to acquire the criterion material if immediate feedback had been provided. This is a particularly imposing hypothesis in light of the extreme differences between the mean exposure times of M and SP (Table 16) under the design of no immediate feedback.

# CHAPTER V

## SUMMARY AND IMPLICATIONS

### Summary

This investigation proposed to study the effects of different rates of presentation of verbal concept material on the subjects' behavior. The Bugelski total time hypothesis was tested at the verbal concept level, and determinations were made concerning effective methodology of CAI material development. The attributes of computer-assisted instruction were used to enhance the flexibility of the experiment, which was basically patterned after standard experimental designs of the verbal learning laboratory.

The Bugelski hypothesis stated that there was a fixed amount of time required to learn a fixed amount of material, and this time was independent of the exposure rate of the material. Other areas of investigation included the subject's ability to function in the environment (psychometer index), his retention, his response time during the recall phase, and his total time to complete the experimental task.

Fifty-two medical students were randomly assigned to four groups differentiated by rates of material exposure. The four time groups were defined as "minimum" (M), "comfortable" (C), "excessive" (E), and "self-paced" (SP). The experimental material concerned vitamin functions and requirements germane to the subjects' experience in pediatrics.

Specifically, there were 15 items which were viewed sequentially at a group-defined rate. After a test over each item, those items answered correctly were removed and the abbreviated list re-presented. Criterion was reached when the subject successfully answered all 15 items. The 15-item test was then re-presented to measure his immediate retention.

It was found that the psychometor index was not an appropriate measure of the kinds of behavior evaluated in this research. It did, however, prove to be an interesting base for evaluating the subjects' adjustment to the CAI environment as they became more familiar with it.

The Bugelski total time hypothesis was rejected at the verbal concept level on the basis of significant differences found between the four groups' total exposure times. There was dramatic graphical evidence of differential behavior by SP concerning total exposure time per question. This was accredited to the design which allowed SP to decide which items were more complex and, therefore, required more study time. The hypothesis of no differences between groups concerning immediate retention scores was accepted for all conditions except those between M and E. The difference between these two groups' immediate retention scores suggested overlearning for E and lack of overlearning for M.

The hypothesis which stated that the group responding times were independent of the exposure rates was rejected. The reason for this rejection, the response time differences between M and C, suggested the existence of a pacing phenomenon. The hypothesis of no difference between groups concerning total amount of time on the test was also rejected. This indicated the existence of two different learning methodologies with relatively equal effectiveness.

## Implications

The introductory chapter defined the development of efficient and effective pedagogy as the foremost problem in the progress of computer-assisted instruction. The generalized implications from this research provided insight into these problems which took the form of appropriate mediational levels[1] and allotment of control during the student's interactive session with the computer.

The rejection of the total time hypothesis at the verbal concept level was due to the differences that existed between the various groups concerning their required times to learn a fixed amount of information. Specifically, group M took significantly less time than any other group to reach criterion of the experimental task. Figure 9 is a bar graph comparing the ratio differences between groups M, C, and E concerning the rates of exposure for each group and the time each required to learn the experimental material. While there is not a one-to-one relationship between the ratio differences of exposure rates and time used, the relationship is a positive, direct one.

The graph defines the minimum time group as the most efficient learning group followed by C, then E. The control group (SP) rated between C and E on this efficiency continuum. From an effectiveness point of view as measured by the immediate retention scores, M was the poorest group. The most effective group was E followed by SP, then C. This empiricism indicates that material which is basic to the acquisition of a broad concept may be most efficiently presented by quick, repetitive

[1]It was assumed by the investigator that the control of the exposure rates affected implicit mediation by the subject since rote memorization (implying no mediation) is not possible at the concept level.

exposures at the CRT. The finding of M's poor effectiveness is not critical if the information is immediately used in another context which will facilitate the student's retention (effectiveness). The idea of rigid proctor control of the learning situation is incongruous with the trends of current educational philosophy. It is also impractical to develop a variety of subject matter around this method. Perhaps the most undesirable aspect of this finding is its potential consequence on the students' attitude. If the method of presenting the material at the computer terminal has a negative effect on the attitude of the student, then no amount of learning efficiency can compensate.

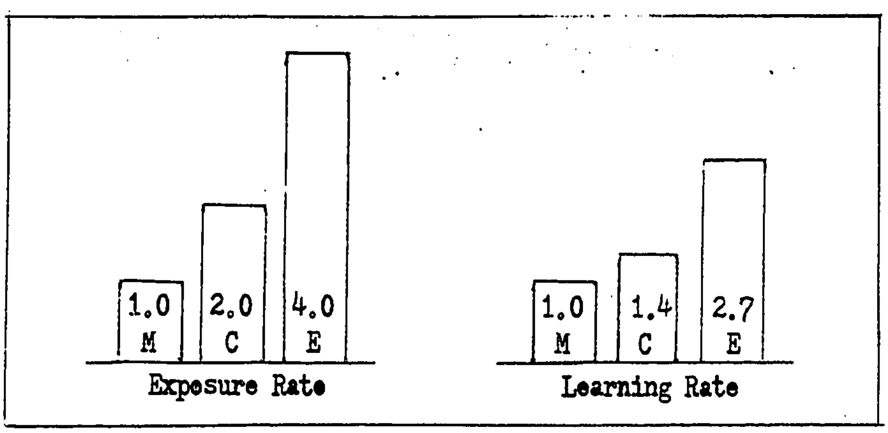



Fig. 9.--A comparison of the differences (in seconds) in exposure rates and time required to learn (Groups M, C, E)

The subjects in C were operating under the poorest overall learning conditions. The mediation control was relatively high with the viewing time only twice that of the base reading time. The subjects were given time to begin mediation, but the process was likely being interrupted by an item change on the CRT. The recognizable pace set by the computer during the study phase had a debilitating effect on C's efficiency while responding during the recall phase. Their effectiveness was not significantly different from E's, but the average items recalled on the immediate retention test was numerically smaller than any other group except M (Table 19). It may be concluded that the data obtained from this group contraindicates the development of any comfortably paced material presentation in CAI.

E's effectiveness in recalling the experimental material on the immediate retention test was shown to be superior in Tables 19 and 22. This suggested overlearning by E and was discussed in Chapter 4. The general characteristics of the treatment received by E included a minimum control of mediation by the computer and excessive viewing times. While these characteristics facilitated their effectiveness, they were shown to be the least efficient group.

SP used the least amount of time attending to the total task (Table 32). The mediation control by the computer was minimal with the viewing time under the subject's control. This combination provided the soundest learning situation with respect to the complete task (acquisition, immediate recall, and retention).

The data from this research supports the following statements: For efficiency during the acquisition of information using CAI, the student may be subjected to a situation in which both mediational processes and types of presentation are completely controlled by the computer. Effectiveness is not facilitated by this method. For the most effective CAI presentation, the student may be subjected to a situation of minimal computer control of mediational processes and maximal control of the amount of time the information is available. This method does not enhance efficiency. The significant conclusion of the study was that the best combination of time efficiency and material retention in the CAI environment occurred when the subject was allowed to control his own exposure rate.

## LIST OF REFERENCES


Baumeister, A. A., and Hawkins, W. F. 1966. Stimulus-response durations in paired-associate learning. Psychon. Science 4: 167.

Bilodeau, E. A., and Bilodeau, I. M. 1958. Variation of temporal intervals among critical events in five studies of knowledge of results. J. exp. Psychol. 55: 603-612.

Bousfield, W. A., Sedgewick, C. H. W., and Cohen, B. H. 1954. Certain temporal characteristics of the recall of verbal associates. Amer. J. of Psychol. 67: 111-118.

Braun, H. W., and Heymann, S. P. 1958. Meaningfulness of material, distribution of practice, and serial position curves. J. exp. Psychol. 56: 146-150.

Bugelski, B. R. 1962. Presentation time, total time, and mediation in paired-associate learning. J. exp. Psychol. 63: 409-412.

Bugelski, B. R., and Rickwood, J. 1963. Presentation time, total time, and mediation in paired-associate learning: Self-pacing. J. exp. Psychol. 65: 616-617.

Bunderson, C. V. March-June, 1966. Quarterly Progress Report of the Laboratory for CAI. Austin: University of Texas.

Carroll, J. B., and Burke, M. 1965. Parameters of paired-associate verbal learning: Length of list, meaningfulness, rate of presentation, and ability. J. exp. Psychol. 69: 543-553.

Ceraso, John. October, 1967. The interference theory of forgetting. Scientific American 217: 117-124.

Deese, J. 1957. Serial organization in the recall of disconnected items. Psychol. Reports 3: 577-582.

Deese, James, and Kaufman, Roger. 1957. Serial effects in recall of unorganized and sequentially organized verbal material. J. exp. Psychol. 54: 180-187.

Entwisle, G., et al. 1963. The use of the digital computer as a teaching machine. J. Med. Educ. 48: 803-812.

Feurzeig, W., et al. 1964. Computer-aided teaching in medical diagnosis. J. Med. Educ. 39: 746-754.

Glucksberg, S., and Laughery, K. 1965. Sequential memory as a function of total time of information exposure and availability of information-processing strategies. Proceedings of the 73rd Annual Convention of the American Psychological Association. pp. 79-80.

Goss, A. E., Morgan, C. H., and Golin, S. J. 1959. Paired-associates learned as a function of percentage of occurrence of response members (reinforcement). J. exp. Psychol. 57: 96-104.

Hovland, C. E. 1949. Experimental studies in rote-learning theory: VIII. Distributed practice of paired-associates with varying rates of presentation. J. exp. Psychol. 39: 714-718.

Johnson, N. F. 1964. The functional relationship between amount learned and frequency vs. rate vs. total time of exposure of verbal materials. J. verb. Learn. verb. Behav. 3: 502-504.

Kausler, Donald H., ed. 1966. Readings in Verbal Learning: Contemporary Theory and Research. New York: John Wiley & Sons, Inc.

Keppel, G., and Rehula, R. J. 1965. Rate of presentation in serial learning. J. exp. Psychol. 69: 121-125.

Murdock, B. B., Jr. 1965. A test of the "limited" capacity hypothesis. J. exp. Psychol. 69: 237-240.

Newman, S. E. 1964. Effect of pairing time and test time on performance during and after paired-associate training. Amer. J. Psychol. 77: 634-637.

Noble, Clyde E. 1952. An analysis of meaning. Psychol. Rev. 59: 421-430.

Nodine, C. F. 1965. Stimulus durations and stimulus characteristics in paired-associate learning. J. exp. Psychol. 69: 534-536.

Peterson, Margaret Jean, Colavita, Francis B., Sheahan, Drexel B. III, and Blattner, Keith C. 1964. Verbal mediating chains and response availability as a function of the acquisition paradigm. J. verb. Learn. verb. Behav. 3: 11-18.

Postman, L., and Goggin, J. 1964. Whole versus part learning of serial lists as a function of meaningfulness and intralist similarity. J. exp. Psychol. 68: 140-150.

Postman, L., and Goggin, J. 1966. Whole versus part learning of paired-associate lists. J. exp. Psychol. 71: 867-877.

Rock, I. 1957. The role of repetition in associative learning. Amer. J. Psychol. 70: 186-193.

Steel, Robert G. D., and Torrie, James H. 1960. Principles and Procedures of Statistics. New York: McGraw-Hill Book Company, Inc.

Underwood, B. J. 1963. Stimulus selection in verbal learning. Verbal Behavior and Learning. Edited by C. N. Cofer and B. S. Musgrave. New York: McGraw-Hill Book Company, Inc.

Underwood, B. J., and Keppel, G. 1963. Bidirectional paired-associate learning. Amer. J. Psychol. 76: 470-474.

Underwood, B. J., Runquist, Willard N., and Schulz, Rudolph W. 1959. Response learning in paired-associate lists as a function of intralist similarity. J. exp. Psychol. 58: 70-78.

Underwood, B. J., and Schulz, Rudolph W. 1960. Meaningfulness and Verbal Learning. Philadelphia: Lippincott.

Wilcoxon, H. C., Wilson, W. R., and Wise, D. A. 1961. Paired-associate learning as a function of percentage of occurrence of response members and other factors. J. exp. Psychol. 61: 283-289.

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