MODE OF PRESENTATION, PACING, KNOWLEDGE OF RESULTS, AND INTELLECTUAL LEVEL IN AUTOMATED INSTRUCTION

by

GARY W. EVANS

Bachelor of Science
Kansas State College at Pittsburg
Pittsburg, Kansas
1957

Master of Science
Kansas State College at Pittsburg
Pittsburg, Kansas
1959

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
May, 1963

MODE OF PRESENTATION, PACING, KNOWLEDGE OF RESULTS, AND INTELLECTUAL LEVEL IN AUTOMATED INSTRUCTION

Thesis Approved:

William W. Rambo
Thesis Adviser

R. W. Kraft

Roy Gladstone

Henry A. Cross

Thomas

Robert Nordquist
Dean of the Graduate School

ACKNOWLEDGEMENTS

I would like to express my thanks and acknowledge the role of several individuals in the completion of this research project.

I am especially grateful to Dr. William W. Rambo, Chairman of the doctoral committee, for his constructive criticism and generous investment of time. I am also indebted to Drs. Gladstone, Scofield, and Roman who served on the doctoral committee. Special thanks are extended to Dr. Edgar Haverland for information regarding material which was unpublished or of limited accessibility. Finally, I would like to thank my wife, Carol, who, as tabulator and typist, played an important role in the completion of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

THE PROBLEM

Among recent approaches to the problem of increasing instructional efficiency, those of automating the instructional process and the use of "teaching machines" appear most fruitful. Lumsdaine (1959) states "automated instructional methods may be considered generally to comprehend any means, devices of materials, whereby teacher or tutor functions are replaced, or provided, by a wholly or partially automated sequence of instructional segments that is prepared in advance and is capable of instructing effectively when presented without direct intervention of modification by a teacher."

Although there is rather general agreement among psychologists that automated instruction is an efficient method of instruction there are many dimensions involved in the process of automated instruction itself upon which there is a divergence of opinion and upon which experimental evidence is either completely lacking or controversial.

The purpose of this study is to investigate experimentally some dimensions of automated instruction upon which there remains some controversy. More specifically, the purpose of this study is to investigate the influence of four factors involved in the process of automated instruction upon performance on a subsequent examination covering the instructional material and upon the frequency of errors committed (error rate) on the program itself. The following factors were investigated:

1. Mode of presentation of subject matter material. The relative advantages of teaching machines and programmed textbooks have been

discussed at some length (Skinner, 1954; Corrigan, 1959; Hively, 1959; Homme and Glaser, 1959). The determination of just what functions the machine should provide beyond such obvious ones as automatic recording and prevention of cheating by the subject is not, as yet, certain. Whether or not the machines have intrinsic motivational properties which cannot be matched by programmed textbooks has not been satisfactorily resolved.

2. Type of pacing utilized. Concerning self-pacing versus experimenter-pacing, the general belief has been that self-pacing is more efficient. Galanter (1959), however, believes that a completely subject controlled pace is wasteful because of the occasional long response latencies invariably present under this condition. Although some experimental attention has been directed at this problem the most effective type of experimenter-pacing has probably not been used in these studies (Briggs, Plashinski, and Jones, 1959; Silverman and Alter, 1961). This problem has important implications concerning the feasibility of group instruction via automated teaching devices.

3. Schedule of reinforcement. To date, the effects of various schedules of reinforcement in automated teaching devices have not been investigated. Galanter (1959) suggests that a partial reinforcement schedule might enhance the inherent interest in the subject matter, as it seems to in animal experiments. On the other hand, as Zeaman (1959) points out, the paradigm which we find to be present in the case of the teaching machine seems to contain characteristics of the free operant model, the controlled operant model and the classical conditioning model; therefore, it is difficult to predict on the basis of previous work just what effect partial reinforcement will have on learning via automated

instructional principles. Amsel (1960) reports that Skinner feels the most advantageous reinforcement schedule for automated instructional devices would be a variable-ratio schedule. This schedule would seem especially well suited to maintain a high rate of machine operation. Amsel, himself, believes that partial reinforcement probably should not be employed as a condition of training in automated instruction because the same response is not recurring as in the case of building up rate of responding or persistence of responding of a simple instrumental response. Amsel feels, however, that partial reinforcement schedules may be important in training which employs automatic devices in two respects: (a) to maintain an optimal level of device operation in the acquisition of the shaped behavior, and (b) to increase the persistence of behavior once it has been shaped. Carr (1960) also distinguishes between the events which might get reinforced, i.e. responding, per se, and responding correctly.

4. Individual differences in regard to measured intelligence. The suggestion has been made that it may be possible to refine programs to the extent that they are free of intelligence and prior achievement effects. Indications at the present time are that, if this is possible, programs have not yet been refined to this point; however, no systematic attempt has been made to assess the effectiveness of automated instruction on subjects with varying degrees of measured mental ability.

The following hypotheses were advanced concerning performance on the criterion examination:

1. Groups receiving instruction via teaching machines will score higher on the criterion examination than groups receiving instruction

via programmed textbooks due to intrinsic motivational properties of machines postulated by Skinner (1954) and Hively (1959).

2. Groups within which the subjects set their own rate of progress (pace) will score higher on the criterion examination than groups whose pace is set by the experimenter due to the fact that each individual is best able to determine his optimal pace and any other pace is less efficient.

3. Groups which receive only partial knowledge of results on a variable-ratio schedule will score higher on the criterion examination than groups receiving knowledge of results after each response. This schedule might operate in such a manner as to enhance the inherent interest in the subject matter (Galanter, 1959).

4. Performance on the criterion test will be positively related to the measured intelligence of the groups. This prediction is based on indications, in other studies, that programs have not yet been refined to the point that they are entirely free of the effects of intelligence and prior achievement.

The following hypotheses were advanced concerning error rate on the program itself:

1. Groups receiving instruction via teaching machines will have lower error rates than groups receiving instruction via programmed textbooks.

2. Groups within which the subjects set their own pace will have lower error rates than groups whose pace is set by the experimenter.

3. Groups which receive only partial knowledge of results will have a higher error rate than groups receiving knowledge of results after

each response because an error made on an unreinforced trial will tend to be repeated on the next trial.

4. The measured intelligence level of a group will be inversely related to its error rate.

Excluding the knowledge of results conditions, these predictions are based on the same factors as were the predictions concerning performance on the criterion examination.

A last hypothesis is that the experimental treatment effects are independent of one another (i.e. they do not interact). Since these factors have not been studied in combination before, little is known about the probability of various interactions.

CHAPTER II

REVIEW OF THE LITERATURE

The concept of an automated teaching device first appeared in S. L. Pressey's (1926) description of a device which "gives tests and scores and teaches." His device was a simple apparatus, resembling a typewriter, utilizing multiple choice questions, on which four keys represented four alternative answers. When the correct key was depressed the machine rolled up a new question, thus both active student response and immediate knowledge of results were incorporated.

Later Pressey and his followers developed several more refined types of apparatus (Pressey, 1927; Pressey, 1932; Angell and Troyer, 1948). Angell and Troyer's device was a punchboard-like piece of apparatus with five perforations for responses by students. When the student chose the correct alternative it was indicated by the appearance of color or other indicia. This effect was achieved by placing a "key" under the answer sheet and inserting both the key and the answer sheet between the front and back cover of the punchboard.

Little (1934) investigated the effect of Pressey's testing and drill machine on final examination scores in an educational psychology course at Ohio State University. Four groups took twelve 30 item tests over units of the course and their scores were tabulated. After each test the drill machine was reset and the student repeated the test until he

could complete it without error. Six control groups took the tests by marking an answer sheet, which was scored that night and returned the next day. The experimental groups' scores were superior to the control groups' scores on both the objective and essay parts of the final; however, the difference was much less on the essay questions.

Jensen (1949) studied the effects of the Angell and Troyer punchboard on scholarship in an educational psychology course at Ohio State University. Two groups of twelve students were chosen on the basis of general ability, academic record, etc. for an accelerated course. These two groups had no class meetings. Instead, each afternoon a laboratory room with the various materials used in the course, a punchboard apparatus, and two or more practice tests on each reading assignment, was open to them. For a final grade, based on a case study, midterm examination and final examination, 54 per cent of the experimental group received A's as compared to 10 per cent of the control group. It should be remembered, however, that the experimental group was composed of superior students.

The next significant advance in the development of automated instructional devices must be attributed largely to the efforts of B. F. Skinner. Skinner (1954) described a teaching machine which, in common with Pressey's, permits self-paced, immediately reinforced practice. An important difference was in the nature of the material inserted in the machine. It is the purpose of this instructional material to present the student with a situation in which the to-be-learned behavior (or knowledge) is acquired through successive approximation of the goal by steps of such small size that correct responding is virtually assured. This, according to Skinner, insures the understanding of all material covered

in a program and, by generating a high percentage of correct responses, minimizes the aversive consequences of failure.

Another major difference between Skinner's and Pressey's approaches was that Skinner expressed a strong preference for devices which required that the student respond by constructing his own answer rather than by recognizing and selecting one of several proferred choices. Since Skinner has taken his position there have been some studies reported concerning both the step size and constructed versus recognition response aspects of automated instruction.

Homme and Glaser (1959) have conducted a study in which four programs with identical content, consisting of 30, 41, 51, and 67 items respectively, were constructed on elementary number theory. Programs consisting of smaller steps (a larger number of items) were associated with better immediate test performance, better retention, and fewer errors on the program.

Coulson and Silberman (1959) using two programs of identical content, consisting of 56 and 104 items respectively have also investigated the importance of step size. Using Harvard's psychological laboratory program as subject matter, they compared two groups of introductory psychology students at Santa Monica on time required to complete all training items without error and comprehension of concepts as measured by a 37 item test (20 fill-in and 17 multiple choice). Small item steps required significantly more training time but also yielded significantly higher test scores than large item steps on the constructed-response criterion subtest. No significant differences were found between experimental groups on the multiple-choice criterion test or on the total criterion test.

Frye (1960) has investigated the relative efficiency of constructed responses and multiple choice responses in the teaching of Spanish words and phrases to ninth grade students in a beginning course in Spanish. Multiple choice and constructed response modes were compared under three conditions:

1. Both groups worked to a criterion of mastery.

2. Both groups worked for a predetermined period of time.

3. Words and phrases were presented to both response groups simultaneously by means of flash cards.

A post-test, consisting of eight multiple choice and eight fill-in items, was given immediately after training by all conditions. A delayed post-test was given two days later. Both immediate and delayed post-tests consistently favored the constructed-response training groups under all conditions. The multiple-choice parts of all post-tests were all very near the maximum, so that all the differences between the groups were seen on the fill-in sections of the post-test.

The studies which have been conducted to date would appear to support Skinner's position (1958) as opposed to Pressey's (1959). That is, shorter steps seem to be superior to longer ones and constructed responses seem to result in more learning than multiple-choice responses.(at least when the post-test includes constructed response items).

Several investigations have been conducted, using Skinner's principles, comparing teaching machine instruction with more conventional types of instruction. Using simple, typewriter-like, teaching machines Porter (1959) investigated the feasibility of their use in teaching second and sixth grade spelling. Twenty-two of the normal 34 weeks of spelling

instruction were given on both the second and sixth grade levels via teaching machines, and student achievement compared to control groups taught in a more conventional manner. On both the second and sixth grade levels spelling achievement, as measured by standardized achievement tests, was significantly superior for the experimental groups.

Ferster and Sapon (1959) have applied Skinner's teaching machine principles to the teaching of German composition. The materials used were sheets of paper with a mask which permitted the exposure of one line at a time. The student exposed one line which presented a German sentence. He then responded, on a piece of scratch paper, with what he believed to be the English translation of the sentence. He then exposed the next line which contained the correct English translation. The subjects (mostly volunteer graduate students with no prior experience in German) were allowed to study where and when they desired; however, they were given instructions to keep an exact record of the time they spent working. Subsequent examinations revealed that, in a mean time of 47.5 hours, the six subjects who finished the course learned an amount of German equivalent to that presented in a first semester course.

Skinner and Holland (1958) and Holland (1959) have described the results of a psychology program used in an introductory psychology course. Machine instruction was made available (1400 frames covering about 200 pages in the textbook) to students enrolled in Harvard's introductory psychology course. An analysis of performance on the final examination showed that items on which the machine could have been expected to help were much more frequently answered correctly than were items for which the machine work was irrelevant. A questionnaire given

following machine instruction showed that 78 per cent of the students checked one of the two following items: (a) "I learned much more on the machine." (32 per cent); and (b) "I learned somewhat more on the machine." (46 per cent).

Findings such as these indicate that automated instruction is an efficient means of instruction but they do not delineate the factors which make it so.

Related Research. Homme and Glaser (1959) have adapted Skinner-type programs to book form in order to have what they call a programmed textbook. There are several frames on each page with the answer to each frame appearing on the reverse side of the page. All frames at the top of the page are gone through first, at which point the student returns to the first of the book and then goes through the second frame on the pages, etc. Although the programmed textbook has incorporated the major principles of the teaching machine, there are certain problems associated with its use. These problems are that cheating cannot be controlled (Corrigan, 1959); there is no way of dropping out items responded to correctly (Corrigan, 1959); and finally, the book does not have the supposed motivational attributes of the machine presentation (Skinner, 1954; Hively, 1959). Homme and Glaser (1959) maintain that cheating possibilities may not be as important as believed because:

1. Cheating may act as a prompt for learning and may not be damaging at all.

2. Possibly through more adequate programming cheating will occur so infrequently that it will not constitute a problem of any magnitude. When the correct response tendency is weak cheating is more apt to occur.

A weak response tendency is the result of inadequate programming. Adequate programming might also eliminate the necessity of dropping out items which have been responded to correctly as adequate programming should, according to Skinner and his followers, allow the student to progress through the program without any errors. Actually only the most expensive machines make provision for the dropping out of items which have been responded to correctly so this particular criticism cannot be construed as especially damaging to programmed textbooks.

In any event, results reported by Homme and Glaser (1959) indicate that the programmed textbook is superior to a conventional textbook when the material covered is identical. Three related experiments were reported. Two of these experiments utilized a sequence of items written directly from a portion of a standard statistics text, and the third utilized a sequence presenting the fundamentals of music reading. Results showed that, in general, the subjects using the programmed textbook made higher achievement scores than did subjects receiving conventional presentation of the same materials.

In order to determine just what function the machine serves beyond such obvious ones as automatic recording and prevention of cheating by the student, Silverman and Alter (1961) compared performance of subjects who had received instruction on psychology and binary numbers programs via either an elaborate machine, a crude machine or a programmed booklet. The differences between devices were not significant. This experiment provides no support for the contention that machines have motivational properties not provided by programmed textbooks.

Descriptions of automated teaching devices invariably include some reference to the fact that the subjects are able to proceed at their own

rate. Unquestionably this has obvious advantages, but it does not necessarily follow that the efficiency of automated teaching depends at all or in part on this self-pacing feature. Galanter (1959) believes that either a completely subject-controlled or experimenter-controlled pace is wasteful. Long pauses between some responses by students is inefficient and the delay between the response and the presentation of reinforcement, as sometimes occurs in experimenter paced instruction, is likewise inefficient. Galanter believes the solution might be to compute an average time to answer and then slowly decrease the time allowed to respond.

Briggs, Plashinski and Jones (1959) found no difference between subjects who were self-paced and those who were machine-paced when the learning task was a 20 item paired-associate list. It should be noted, however, that paired-associate lists differ from the continuous discourse material used in teaching machine programs.

Silverman and Alter (1961) compared programmed learning under paced and unpaced conditions. Three groups were given automated instruction over an 87-item unit dealing with basic electricity. Two groups were paced according to different schedules and the third group was unpaced. No significant differences were found among the three experimental groups. One difficulty with this study was that they had no empirical basis for the determination of item times, consequently the most efficient experimenter-pace probably wasn't used.

To date no experiments have been reported which deal with the effect of partial reinforcement on programmed instruction. The controversy over the probable outcome of such an investigation was reported in Chapter I.

Skinner (1958) points out that different programs may have to be constructed for students at different intelligence levels. Studies which have investigated levels of intelligence in relation to methods of instruction have found there was no relationship between intelligence scores and achievement in groups receiving automated instruction (Porter, 1959) and that although groups receiving automated instruction made higher achievement scores than groups receiving conventional instruction, they exhibited less variability of performance than groups receiving conventional instruction (Homme and Glaser, 1959). This might indicate that automatic instructional devices, with their small steps, might benefit slow learners while holding the more rapid learners back. This would indicate the possible desirability of introducing larger steps for the more able students (Pressey, 1959). On the other hand, these results may indicate that relatively complete mastery of the subject matter can be acquired by all intellectual levels thus far investigated. The beneficial effect of small steps found in previously mentioned studies (Homme and Glaser, 1959; Silberman and Coulson, 1959) would seem to support the latter interpretation.

In summary, it can be said that preliminary investigations seem to indicate the superiority of automated instructional methods over conventional instruction in the situations in which the two have been compared (Jensen, 1949; Ferster and Sapon, 1959; Porter, 1959; Skinner and Holland, 1959; and Holland, 1959).

The dispute between Skinner and Pressey regarding optimum step-size and constructed versus multiple-choice responses has received some experimental attention and results seem to favor Skinner's preferences

for small steps (Homme and Glaser, 1959: Coulson and Silberman, 1959) and constructed responses (Frye, 1960).

Other issues associated with the use of automated instructional devices such as the beneficial effects of machines, self-pacing and schedules of knowledge of results have not been extensively studied.

CHAPTER III

METHOD

The Experimental Sample. Fifty-eight students enrolled in the introductory psychology course (Psychology 213) at Oklahoma State University during the fall semester of 1961, were used as subjects in the experiment. Of these, ten were used in order to determine optimal group pacing conditions and the remaining 48 were used in the experiment proper. Subjects had been in the psychology course for about one month, covering some general background in the area of psychology, but had not been given specific instruction in any material closely related to the subject material used in the experiment. None of the subjects had had a prior course in psychology.

Materials. The following materials were used in this investigation:

1. Four Foringer Teaching Machines (standard models).

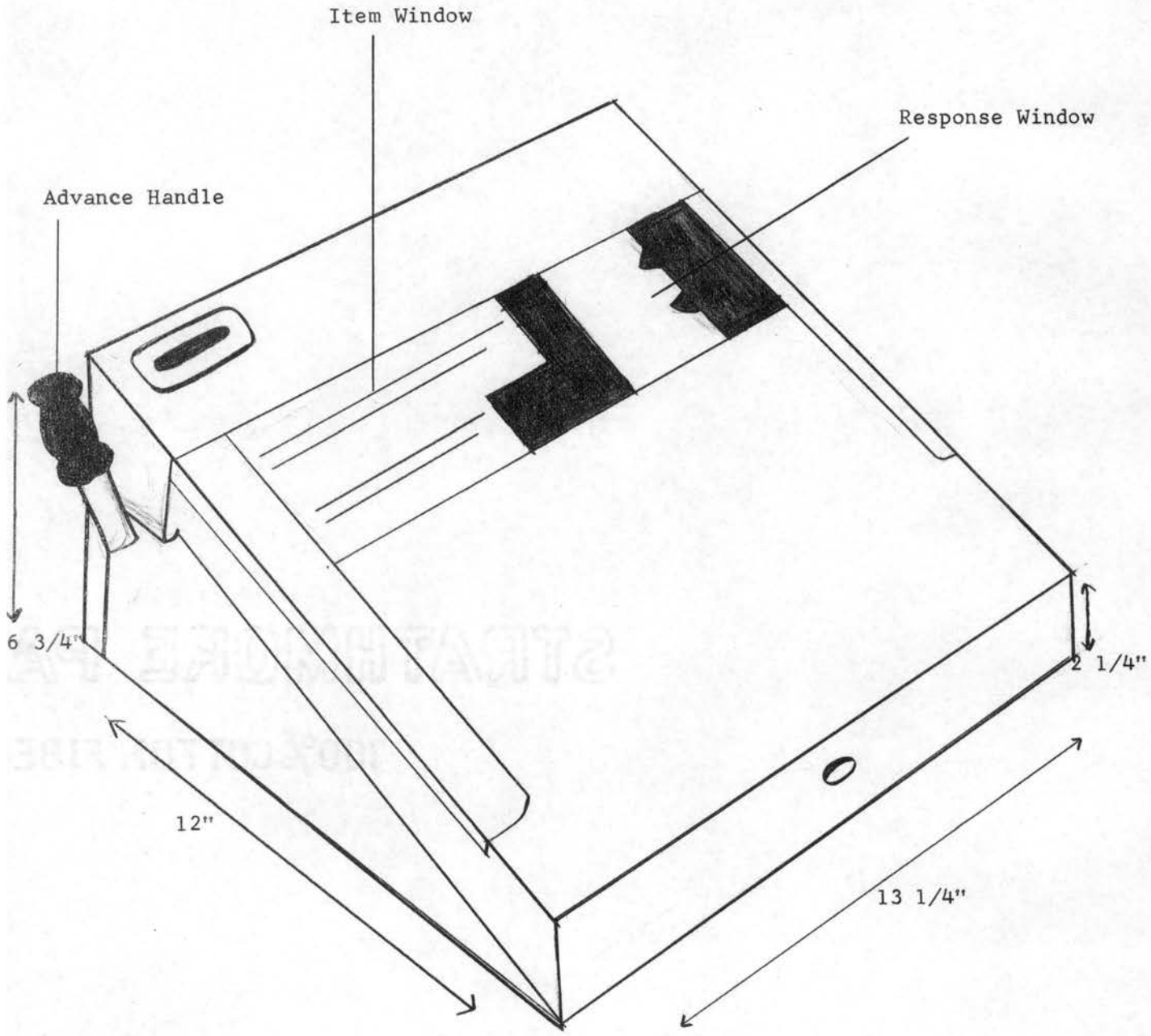
2. Four copies of Holland and Skinner's (1961) programmed textbook, Analysis of Behavior.

3. Ninety-two copies of the Otis Quick-Scoring Mental Ability Tests (Gamma Test).

An idea of the appearance of the Foringer Teaching Machine may be obtained from Illustration I. The machine is shaped much like a box with a sloping surface. It is 12 inches in length and 13 1/4 inches in width. The height at the back of the machine is 6 3/4 inches and 2 1/4

Foringer Teaching Machine

(Standard Model)



inches at the front. The most important features of the machine, in order to understand its operation, are labeled in Illustration I as "Advance Handle", "Item Window" and "Response Window". Items to be responded to are exhibited in the "Item Window". The student writes what he believes to be the correct response on the response tape in the "Response Window". He then pushes the "Advance Handle" attached to the left side of the machine (right side for left-handed subjects) forward. This places tension on a roller inside the machine and the test strip and response tape are advanced, allowing the next item and the correct response to the previous item to be presented in the "Item Window" (the correct response is shown in the upper right portion of the window).

The programmed textbooks used were 6 by 9 inch paperbacked books which differ from other books only in content and the manner in which they are used. The book consists of a series of incomplete statements with words or phrases to be filled in by the student.

The Gamma Test of the Otis Quick-Scoring Mental Abilities Tests is an easily administered and scored group test developed to test students ranging from the ninth grade through college. Norms are based, in part, on 3180 pupils in this educational range.

"The reliability of the Gamma Test was investigated by correlating the odd-numbered and even-numbered items of the test papers of 257 pupils in Grades 10, 11 and 12. The coefficients so found were, respectively, .82, .85, and .73 for the three grades. When corrected by the Spearman-Brown formula, the reliability coefficients for the three grades were found to be .90, .91, and .85, respectively." (Otis, 1939).

Procedure. The experiment consisted of two phases: (a) training for each of the 48 subjects under one of the eight types of automated instruction being compared, and (b) a criterion quiz following his

training session to determine how much the subject learned in the training session.

The training for any single subject consisted of one 50 minute session per day for three consecutive days. The criterion examination was administered two days after the third training session.

Four identical teaching machines and four identical programmed textbooks were used in the experimental training sessions. Each machine and each programmed textbook was operated under one of four experimental conditions. These represented the eight combinations of three of the experimental variables, each variable having two possible levels as follows: (a) mode of presentation (teaching machine versus programmed textbook); (b) mode of pacing (subject-pacing versus experimenter-pacing); and (c) schedule of reinforcement (100 per cent reinforcement versus partial reinforcement).

Assignment of subjects to treatment combinations was accomplished by classifying available subjects into three intellectual levels and then randomly assigning two subjects from each level to each treatment combination.. A schematic representation of the experimental design is presented in Table I.

Scores on the Otis Quick-Scoring Mental Ability Tests (Gamma Test) furnished the basis for the classification of subjects into intellectual levels. The test was administered to 92 introductory psychology students and on the basis of this distribution, 16 subjects whose scores were in the lowest quartile, 16 subjects whose scores fell in the semi-inter-quartile range, and 16 subjects whose scores were in the highest quartile comprised the three intellectual groups which was the fourth

TABLE I

SCHEMATIC REPRESENTATION OF THE EXPERIMENTAL DESIGN

Highest Intellectual Level	Machine Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement
	Programmed Text Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement
Intermediate Intellectual Level	Machine Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement
	Programmed Text Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement
Lowest Intellectual Level	Machine Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement
	Programmed Text Presentation	Self-Paced	100 per cent reinforcement
			Partial reinforcement
		Experimenter-Paced	100 per cent reinforcement
			Partial reinforcement

factor investigated in this study. The distribution of raw scores is presented in Appendix A. Those scoring above 61 were in the highest quartile, those scoring from 52 to 57 inclusively were in the semi-inter-quartile range and those scoring below 50 were in the lowest quartile.

Printed instructions explaining the task to be performed were given to each subject at the beginning of the first training period. After the subject had read the instructions he was given the opportunity to ask questions concerning any aspects of the task which were not clear to him. The instructions given the respective experimental groups are presented in Appendix B.

A description of the three experimental conditions follows:

1. Teaching machine presentation versus programmed textbook presentation. Subjects receiving teaching machine presentation, upon entering the experimental room, were seated before a machine and given instructions converning its operation. Holland and Skinner's program, Analysis of Behavior was reproduced in suitable form so the items and displays could be presented to subjects receiving machine instruction. This was the instructional material used in the experiment.

It should be mentioned that subjects receiving teaching machine instruction wrote their responses on a separate answer sheet rather than on an answer tape which could have been inserted in the machine. The reason for this was that those subjects receiving programmed textbook instruction, to be described shortly, used separate answer sheets and should the results have indicated superiority on the part of the subjects receiving machine instruction it would not have been possible

to interpret this difference as due to motivational properties of the machines since another systematic variation between the two treatment conditions would have been present.

Four copies of Holland and Skinner's (1961) programmed textbook, The Analysis of Behavior, were used by subjects receiving programmed textbook instruction. The program contained in this book consists of a series of incomplete statements. The subject's task was to complete each statement by writing a word or words on a separate answer sheet provided, check his answer by turning the page and looking at a designated location, and then respond to the next item which is adjacent to the correct answer of the previous item. Six items are presented on each page. The subject was instructed to respond to the items at the top of the pages until the program instructed him to go back to the first page. Then he was instructed to respond to the items directly below the items at the top of the page (i.e. items second from the top), etc.

2. Self-pacing versus experimenter-pacing. Under the self-pacing condition subjects were free to move from item to item at their own rate.

Under the experimenter-pacing condition subjects moved from item to item at a predetermined rate. The subject had a specified period of time to make his response(s) to an item. The subject was not allowed to move to the next item until the experimenter indicated that time had expired. At this point he was told to go to the next item, even if he had not completed his response to the previous one.

As the amount of information contained in the various items differ and as the number of responses called for vary from item to item it was

necessary to establish individual item exposure times empirically for the experimenter-paced condition. To accomplish this end, ten randomly selected subjects individually received three 50 minute sessions of instruction over the programmed material prior to the experiment. For each subject individual item completion times were recorded for each item. The median completion time for each item was then computed. These median completion times were used as the item exposure times for experimenter-paced groups in the experiment proper.

The subjects who were experimenter-paced were presented with 334 items (seven sets) of the program during the three 50 minute sessions. The amount of material exposed to the experimenter-paced subjects was determined by summing median completion times of the items which were obtained in the preliminary study. The self-paced subjects, of course, completed a variable number of items. One difficulty with this approach to the problem is that it does not address itself to the question of which type of pacing is superior if time is not important. In other words, either time spent covering a given amount of material or the amount of material covered must be left uncontrolled. The rationale for choosing time as the variable to be controlled in this study was that it seemed reasonable that any beneficial results derived from group pacing would be a function of time saved from the elimination of long response latencies for individual subjects on individual items.

3. One-hundred per cent reinforcement versus partial reinforcement. Under the 100 per cent reinforcement condition subjects received knowledge of results after each response.

Under the partial reinforcement condition subjects received knowledge of results after 50 per cent of their responses on a variable-ratio

schedule. A variable-ratio schedule was chosen since Skinner reportedly feels that this would be the most advantageous schedule for automated instruction (Amsel, 1960). Knowledge of results was not contingent upon a response in the experimenter-paced group.

The two dependent variables used in the experiment were error rate and scores on a criterion test. Failure to respond or an incorrect response was considered an error. Error rate has conventionally been expressed as the percentage of incorrect responses made on the program itself. Alternatives might be to use the total number of errors made on the program or total number of errors with some sort of adjustment for the number of responses made. Total number of errors without adjustment for number of responses doesn't seem to make much sense as a measure when subjects make a variable number of responses and when the same number of responses are made by all subjects the results would be the same as when percentages were used. Actually, percentages are measures of the total number of errors adjusted for the total number of responses. Percentages were used as the error rate measure in this study because they are conventionally used and because they seem to make as much sense as any other measure.

The criterion test consisted of 40 constructed response (fill-in) items and is included in its entirety in Appendix C. All of the questions were based upon the first 486 items (eleven sets) of Holland and Skinner's program. The inclusion of test items from this range of material may have benefited the experimenter-paced subjects since each of these subjects completed 334 items while only four subjects under the self-paced condition completed as many as 334. Few of the questions included on the criterion examination were duplicates of instructional

items, though many of the same words could be used as correct responses. An attempt was made to include questions on the criterion test which covered important concepts spaced rather evenly throughout the material covered. To compute a split-half reliability coefficient the total test was split into two forms by randomly assigning 20 items to each prior to administering the test. Both forms were given to each subject but the form given first was alternated. The two forms correlated .816 which boosted by the Spearman-Brown Prophecy formula reached .90 for the estimate of reliability for the complete 40 item test.

Treatment of the Data. A 2 x 2 x 2 x 3 factorial arrangement of teatments was used as the experimental design and the statistical analyses of the error rate and criterion test data respectively was performed by means of the corresponding analysis of variance.

Error rate measures are expressed in terms of percentages and, since percentages tend to be distributed rectangularily, an arc-sine transformation was performed on these measures prior to the statistical analysis of the data. This transformation is commonly used when the data is expressed in terms of proportions in order to normalize and equalize variances of distributions. (Walker and Lev, 1953).

The F-test of analysis of variance assumes that the treatment variances are equal and the treatment distributions are normally distributed in the population. Due to the fact that only two observations per cell were available these assumptions were not tested for either the error rate data or the criterion test data. This small number of observations makes a test of normality impossible and no meaningful statistic is available for testing equality of variances when the number of observations

is as small as two. The use of a statistical test, the accuracy of which is dependent upon fulfilling certain assumptions is of course undesirable, however, Box (1954) has shown that a slight departure from the assumptions usually will not cause serious error in the F-test.

A preliminary F-test was made for each of the two criteria in order to determine if there were any overall treatment effects. Following this test the treatment sum of squares was partitioned, by means of two-way tables, into components due to mode of presentation, type of pacing, reinforcement schedule, intellectual levels, and the various interactions. A subsequent orthogonal contrast was made to test the intellectual levels treatment effect for linearity. A difference was considered statistically significant if it reached or exceeded the .05 level of significance.

CHAPTER IV

RESULTS

The raw data used in the analysis is presented in Appendix D.

The observed mean criterion test scores for the different experimental conditions and the mean scores for various combinations of the experimental conditions are presented in Table 2. The results of the statistical analysis of the data are presented in Table 3.

Table 3 shows the only significant treatment effect found in the criterion test score analyses was that for intellectual levels. Table 2 shows that the group composing the highest of the three intellectual levels had the highest mean score on the criterion examination while the group composing the lowest intellectual level had the lowest mean score on the examination. A subsequent analysis found the F-value for linearity to be significant while the non-linear F was not significant. This indicates that intellectual level was linearly related to scores on the criterion test.

Other comparisons among the experimental conditions resulted in the following findings:

1. Table 2 shows that the mean criterion test score for subjects receiving instruction via programmed textbooks was slightly higher than the mean score for subjects receiving instruction via teaching machines; however, as Table 3 reveals, this difference did not approach statistical significance.

TABLE 2

OBSERVED MEAN SCORES ON CRITERION TEST

Groups	TM[1]	PT[2]	100% R[3]	50% R[4]	SP[5]	EP[6]	HIL[7]	IIL[8]	LIL[9]	Grand Mean
Machine	-----	-----	20.00	16.83	16.58	20.25	23.00	18.00	14.25	18.42
Programmed textbook	-----	-----	18.42	18.67	18.08	19.00	23.00	17.38	15.25	18.54
100% reinforcement	20.00	18.42	-----	-----	18.92	19.50	22.75	20.63	14.25	19.21
50% reinforcement	16.83	18.67	-----	-----	15.75	19.75	23.25	14.75	15.25	17.75
Self-paced	16.58	18.08	18.92	15.75	-----	-----	22.50	16.25	13.25	17.33
Experimenter-paced	20.25	19.00	19.50	19.75	-----	-----	23.50	19.13	16.25	19.63
Highest Intellectual level	23.00	23.00	22.75	23.25	22.50	23.50	-----	-----	-----	23.00
Intermediate intellectual level	18.00	17.38	20.63	14.75	16.25	19.13	-----	-----	-----	17.69
Lowest intellectual level	14.25	15.25	14.25	15.25	13.25	16.25	-----	-----	-----	14.75
Grand mean	18.42	18.54	19.21	17.75	17.33	19.63	23.00	17.69	14.75	18.49

1. TM = Teaching machine
2. PTB = Programmed textbook
3. 100% R = 100% reinforcement
4. 50% R = 50% reinforcement
5. SP = Self paced
6. EP = Experimenter paced
7. HIL = Highest intellectual level
8. IIL = Intermediate intellectual level
9. LIL = Lowest intellectual level

TABLE 3

ANALYSIS OF VARIANCE OF
CRITERION TEST SCORES

Source of Variation	Adjusted Sum of Squares	Degrees of Freedom	Mean Square	F
Experimental Treatment Combinations	1142.476	23	49.673	1.765
A. Mode of Presentation	.188	1	.188	-----
B. Reinforcement Schedule	25.521	1	25.521	-----
C. Type of Pacing	63.021	1	63.021	2.239
D. Intellectual Levels	559.542	2	279.771	9.940*
Linear	544.501	1	544.501	19.346*
Non-linear	15.041	1	15.041	-----
AB	35.020	1	35.020	1.244
AC	22.688	1	22.688	-----
AD	5.374	2	2.687	-----
BC	35.020	1	35.020	1.244
BD	117.541	2	58.771	2.088
CD	10.041	2	5.021	-----
ABC	46.021	1	46.021	1.635
ABD	53.042	2	26.521	-----
ACD	81.374	2	40.687	1.446
BCD	73.042	2	36.521	1.298
ABCD	15.041	2	7.521	-----
Within Cells	675.504	24	28.146	
Total	1817.980	47		

*.05 level of significance.

2. Subjects receiving knowledge of results after each response (one-hundred per cent reinforcement) had a higher mean score on the criterion examination than subjects receiving only partial knowledge of results. Again, however, the difference was not statistically significant.

3. The experimenter-paced group had a higher mean score on the criterion examination than the self-paced group but the difference was not statistically significant.

A chi-square analysis was performed to test for a difference in amount of material covered by subjects under the self-paced condition. Each subject under the experimenter-paced condition was exposed to 334 programmed items of programmed material while only four of the 24 self-paced subjects were exposed to this many items. This yielded a chi-square value of 10.67, with one degree of freedom, which is highly significant.

Since a difference was found in amount of material exposed to the respective groups, a Pearson correlation coefficient was computed between items completed and criterion test scores for the self-paced subjects. The Pearson r was found to be .805. Due to this rather high correlation coefficient the criterion test mean was adjusted through a regression equation to 334 programmed items for the self-paced group (the number of items exposed to the experimenter-paced subjects). The adjusted criterion test mean for the self-paced group was 23.66 as compared to a mean of 19.63 for the experimenter-paced group. An analysis of variance F-test was computed to test for a statistically significant difference between these means. The resulting F-value was 6.89, with 1 and 24 degrees of freedom, which is statistically significant.

4. None of the interaction components were found to be statistically significant.

The observed mean error rates, expressed in percentages, for the different experimental conditions and the mean scores for various combinations of experimental conditions are presented in Table 4. The results of the statistical analysis of the transformed data are presented in Table 5.

The F-value for type of pacing conditions was statistically significant with the self-paced group having a lower mean error rate than the experimenter-paced group.

The F-value for intellectual levels was significant. A subsequent analysis found the F-value for the linearity component to be significant while the F-value for the non-linear component was not significant. The group composing the highest of the three intellectual levels had the lowest error rate while the group composing the lowest intellectual level had the highest error rate.

The F-values for mode of presentation, schedule of reinforcement and the various interaction components were not statistically significant.

TABLE 4

OBSERVED MEAN ERROR RATE SCORES

Groups	TM[1]	PT[2]	100% R[3]	50% R[4]	SP[5]	EP[6]	HIL[7]	IIL[8]	LIL[9]	Grand Mean
Machine	-----	-----	17.42	20.83	15.67	22.58	9.25	18.35	29.75	19.13
Programmed textbook	-----	-----	24.17	23.90	14.50	32.75	15.88	22.88	32.13	23.63
100% reinforcement	17.42	24.17	-----	-----	15.83	25.75	14.00	14.75	33.63	20.79
50% reinforcement	20.83	23.90	-----	-----	14.33	29.59	11.13	26.50	28.25	21.96
Self-paced	15.67	14.50	15.83	14.33	-----	-----	5.00	15.87	24.38	15.08
Experimenter-paced	22.58	32.75	25.75	29.59	-----	-----	20.13	25.38	37.50	27.67
Highest Intellectual level	9.25	15.88	14.00	11.13	5.00	20.13	-----	-----	-----	12.56
Intermediate Intellectual level	18.38	22.88	14.75	26.50	15.87	25.38	-----	-----	-----	20.63
Lowest intellectual level	29.75	32.13	33.63	28.25	24.38	37.50	-----	-----	-----	30.94
Grand mean	19.13	23.63	20.79	21.96	15.08	27.67	12.56	20.63	30.94	21.38

1. TM = Teaching machine
2. PTB = Programmed textbook
3. 100% R = 100% reinforcement
4. 50% R = 50% reinforcement
5. SP = Self-paced
6. EP = Experimenter-paced
7. HIL = Highest intellectual level
8. IIL = Intermediate intellectual level
9. LIL = Lowest intellectual level

TABLE 5

ANALYSIS OF VARIANCE OF TRANSFORMED ERROR RATE DATA

Source of Variation	Adjusted Sum of Squares	Degrees of Freedom	Mean Square	F
Experimental Treatment Combinations	4668.352	23	202.972	3.638*
A. Mode of Presentation	91.300	1	91.300	1.636
B. Reinforcement Schedule	15.870	1	15.870	-----
C. Type of Pacing	1088.707	1	1088.707	19.514*
D. Intellectual Levels	1803.076	2	901.538	16.159*
Linear	1801.500	1	1801.500	32.291*
Non-linear	1.576	1	1.576	-----
AB	17.279	1	17.279	-----
AC	224.467	1	224.467	4.023
AD	40.279	2	20.140	-----
BC	2.803	1	2.803	-----
BD	355.031	2	177.156	3.175
CD	90.797	2	45.399	-----
ABC	83.992	1	83.992	1.506
ABD	260.252	2	130.126	2.332
ACD	94.322	2	47.161	-----
BCD	372.688	2	186.344	3.340
ABCD	127.489	2	63.745	1.143
Within Cells	1338.948	24	55.790	
Total	6007.300	47		

*.05 level of significance.

CHAPTER V

DISCUSSION

The finding that performance on the criterion test did not differ significantly for subjects instructed by programmed textbooks and subjects instructed by teaching machines agrees with the findings reported in a recent review of the research on this topic (Goldstein and Gotkin, 1962). Eight studies were reviewed and none of the eight found a significant difference between the two modes of presentation. These findings provide no support for concluding that machines have unique motivational properties.

Although the difference between teaching machine and programmed textbook conditions was not statistically significant, some people might argue that the lack of statistical significance was due to the insensitivity of the statistical test and/or lack of experimental precision and not due to the absence of a difference in reality. This, of course, may be the case; however, a difference as small as the one found in this experiment (means were 18.54 and 18.42, respectively), based on a fairly large sample of subjects, surely indicates that if teaching machines do have certain unique motivational properties the effect is not large enough to have any practical significance. This position is supported by the finding that error rate means did not differ significantly between the two modes of presentation.

On the basis of the findings in this study and in light of additional findings reported above it seems reasonable to conclude that there is

little reason at this time to be concerned with machine devices for populations comparable to those used in this study. The programmed text appears to be as effective as a machine. While machines may provide certain controls, such as preventing the learner from looking ahead or back, the extent to which these controls are necessary for adult learners is not certain.

This writer does not deny that certain features may be built into machines which make them superior to programmed texts. The findings of this study only indicate that Foringer Teaching Machines and probably other relatively crude machines are not more effective than programmed texts. The findings cannot be generalized to include more complex machines such as those that provide for the dropping out of items responded to correctly or that make the next item presented to the student contingent upon the nature of his response to the previous item. On the other hand, programmed texts have certain positive features which probably cannot be matched by machines. They cost only about as much as a program for a machine, they are portable, thus making home study possible, and they require less space for their storage or use.

The finding that performance on neither of the criteria differed significantly for subjects receiving knowledge of results after each response and those receiving knowledge of results after fifty per cent of the responses on a variable-ratio is in some ways puzzling. The absence of a difference on the criterion test might have been expected on the basis of generalization from studies of partial reinforcement using both humans and animals. The usual finding in these studies is that the reinforcement schedule shows little effect on the acquisition function but a rather large effect in extinction. No extinction measure was used

in this study but even had one been used it is doubtful that an effect would have been found since, as Amsel (1960) has pointed out, in automated instruction the same response is not recurring as in the case of building up rate of responding or persistence of responding of a simple instrumental response.

On the other hand, it is difficult to explain why the error rates of subjects receiving one-hundred per cent reinforcement were not lower than those of subjects receiving knowledge of results only fifty per cent of the time since those subjects on one-hundred per cent schedules should be able to profit more from their incorrect responses and make the corret response the next time it is appropriate. One possible explanation of this surprising finding is the possibility that really good programming makes knowledge of results unnecessary since, ideally, the learner should know the correct response to each item by the time he reaches it (Skinner, 1958). This possibility hardly seems likely in light of the fact that the observed mean error rates for one-hundred per cent knowledge of results groups and fifty per cent knowledge of results groups were 20.79 and 21.96 respectively both of which are considered by most people in the field to be much too high for effective learning. The finding that the reinforcement schedule utilized is not of any critical importance must remain purely empirical until refutation or explanation is offered by further research.

The restriction of experimenter-pacing did not significantly hamper nor did it significantly enhance performance on the criterion test. However, in spite of lack of statistical significance, there was a rather

definite trend which favored the experimenter-paced group. The means were 19.63 and 17.33 for experimenter-paced and self-paced groups respectively. This difference yielded an F-value of 2.239, with 1 and 24 degrees of freedom, which would have been significant had the ten per cent level of confidence been chosen as the point of statistical significance rather than the five per cent level.

On the other hand, subjects operating under the self-paced condition had a significantly lower mean error rate than subjects performing under the experimenter-paced condition. At first thought these findings concerning types of pacing seem contradictory. However, a third finding integrates these two findings. The experimenter-paced group covered a significantly greater amount of material than did the self-paced group. The reason for this difference in the number of items completed was undoubtedly a result of the manner in which individual item times were established for the experimenter-paced condition. The item time used was the median response time of subjects in a preliminary study. Using the median rather than the mean response time for items eliminated the influence of long response latencies for individual items thus making the median response time shorter than the mean response time. When the number of instructional items completed was statistically controlled the mean of the self-paced group on the criterion test was found to be significantly higher than the experimenter-paced group mean.

On the basis of these findings it is obvious that the extent to which pacing affects programmed learning depends both upon the dependent variable used and the speed of pacing employed. If the desire is to cover as much material as possible in a given period of time group pacing

seems very feasible. On the other hand, the error rate analysis and the criterion test analysis following adjustment for differences in the number of programmed items the groups completed indicate self-pacing may result in more complete mastery of the material covered.

This experiment indicates that group pacing is possible and feasible under certain conditions, however, further research should be done to determine optimal pacing rates and methods of pacing in programmed learning situations. Such information may be very useful in training situations where time factors are important and where the material to be learned is composed of relatively independent items.

The clear differences in achievement on the criterion test and on error rates among groups of subjects of different intellectual levels indicate that the program used here was not refined to the point where it was free of intelligence and prior achievement effects. Indeed, it is difficult to imagine a program written for adults which did not make use of their prior reading and attention skills. In light of the relatively high mean error rate for subjects in this experiment it seems likely that the program used in this experiment was too difficult for most effective learning.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The purpose of the present study was to investigate experimentally four dimensions of automated instruction upon which there is some controversy. Forty-eight subjects from introductory psychology classes at Oklahoma State University were given three 50-minute training sessions by means of various types of automated instruction over Holland and Skinner's (1961) program, Analysis of Behavior. The four independent variables were mode of presentation (teaching machine versus programmed text), schedule of reinforcement (one-hundred per cent versus fifty per cent on a variable ratio schedule), type of pacing (self-pacing versus experimenter-pacing) and intelligence level (three levels). Error rate on the program and scores on a criterion test were the dependent variables. An analysis of variance with a factorial arrangement of treatments design was used. Major results were as follows:

1. No significant difference was found between teaching machine and programmed text presentation on either dependent variable.

2. No significant difference was found between schedules of reinforcement conditions on either dependent variable.

3. The mean error rate under the self-paced condition was significantly lower than the mean error rate under the experimenter-paced condition while the mean criterion test score was higher under the

experimenter-paced condition than the self-paced condition although the difference was not statistically significant. When number of instructional items completed was statistically controlled the self-paced group mean was significantly higher than the experimenter-paced group mean on the criterion test.

4. There were significant differences among the three intelligence levels on both dependent variables. Subsequent tests found intellectual level to be linearly related to both error rate and criterion test score (positively to criterion test score and negatively to error rate).

On the basis of these findings the following conclusions were drawn:

1. Teaching machines similar to those used in this study do not have unique motivational properties.

2. The explanation of the role of the reinforcement schedule in automated instruction awaits further research.

3. Self-pacing seems superior to experimenter-pacing when only mastery of the material is considered to be of importance. However, when amount of learning in a specified period of time is the important factor self-pacing is not necessarily superior to experimenter-pacing. The instructional objective and pacing schedule must be considered when attempting to determine whether self-pacing or experimenter-pacing is more efficient.

4. Achievement on the program used in this study, and probably all programs, is not free of intelligence factors and prior achievement effects.

REFERENCES

Amsel, A. Error responses and reinforcement schedules in self-instructional devices. In A. A. Lumsdaine and R. Glaser (Eds.) Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960.

Angell, C. W. & Troyer, M. E. A new self-scoring device for improving instruction. Sch. & Soc., 1948, 67, 84-85.

Beck, J. On some methods of programming. In E. Galahter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

Box, G. E. P. Effects of inequality of variance and of correlation between errors in the two-way classification. Ann. Math. Stat., 1954, 25, 484-498.

Briggs, L. J., Plashinski, D., & Jones, D. L. Self-pacing versus automatic pacing of practice on the subject matter trainer. Unpublished paper, cited in Darby, C. L. An annotated bibliography on the automation of instruction. U. S. Army Air Defense, Human Research Unit, Fort Bliss, Texas, 1959.

Carr, W. J. A functional analysis of self-instructional devices. In A. A. Lumsdaine and R. Glaser (Eds.) Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1960.

Corrigan, R. E. A review of fundamentals, concepts and developments. Automated Teaching Bull., 1959, No. 1, 5-12.

Crowder, N. A. Automatic tutoring by means of intrinsic programming. In E. Galanter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

Evans, J. L., Homme, L. E., & Glaser, R. A preliminary investigation of variation in the properties of verbal learning sequences of the "teaching machine" type. In A. A. Lumsdaine and R. Glaser (Eds.) Teaching Machines and programmed learning. Washington, D. C.: National Education Association, 1960.

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

decomposition classes of $\{\phi_v\}$ are called characteristic subsystems. These characteristic subsystems are ordered by the decomposition property of T2-14. When the system is a parameter model and if the relations are related to some physical interconnection property (e.g., some electrical or mechanical network of components), then the characteristic subsystems indicate the physical interconnection. This is illustrated below:

Example: Consider the relations $\{\phi_i\}$ of $S_{9,4} \in PS$ induced by the Kirchhoff principle of voltages summing to zero in a closed loop.

$$\phi_1 (e_1, e_2, e_3)$$

$$\phi_2 (e_4, e_5, e_6)$$

$$\phi_3 (e_7, e_8, e_9)$$

$$\phi_4 (e_1, e_5, e_9)$$

Now $S_{9,4}$ is decomposable with characteristic subsystems $(\{\phi_1, \phi_2, \phi_3\}, S_{9,3})$ and (ϕ_4, S_{31}). Assume that each ϕ corresponds to traversing a geometrical closed path of components. (One component for each parameter is assumed.) If this is the case, then it must be true that the components of ϕ_4 are connected in a loop which is imbedded in the loops of the system $S_{9,3}$. A network with this property is shown in Figure 3-1.

Porter, D. A. A critical review of a portion of the literature on teach-devices. Harvard Educ. Rev., 1957, 27, 126-147.

Porter, D. A. Some effects of year long teaching machine instruction. In E. Galanter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

Pressey, S. L. A machine for automatic teaching of drill material. Sch. & Soc., 1927, 25, 549-552.

Pressey, S. L. A simple apparatus which gives tests and scores and teaches. Sch. & Soc., 1926, 23, 373-376.

Pressey, S. L. A third and fourth contribution toward the coming "industrial revolution" in education. Sch. & Soc., 1932, 36, 668-672.

Pressey, S. L. Certain major psycho-educational issues appearing in the conference on teaching machines. In E. Galanter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

Pressey, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. J. Psychol., 1950, 29, 417-447.

Silberman, H. F. & Coulson, J. A draft summary of findings in an exploratory teaching machine study. Automated Teaching Bull., 1959, 1, No. 1, 16-18.

Silverman, R. E. & Alter, M. Response mode, pacing and motivational effects in teaching machines. A technical report: NAVTRADEVCEN 507-3. U. S. Naval Training Device Center, Port Washington, L. I., New York, 1961.

Skinner, B. F. Science of learning and art of teaching. Hav. Educ. Rev. 1954, 24 No. 2, 86-97.

Skinner, B. F. Teaching machines. Science, 1954, 128, 969-977.

Skinner, B. F. The programming of verbal knowledge. In E. Galanter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

Skinner, B. F & Holland, J. G. The use of teaching machines in college instruction. In A. A. Lumsdaine and R. Glaser (Eds.) Teaching machines and programmed learning. Washington, D. C.: National Education Association, 1958.

Walker, H. M. & Lev. J. Statistical inference. New York: Henry Holt and Co., 1953.

Zeaman, D. Skinner's theory of teaching machines. In E. Galanter (Ed.) Automatic teaching. New York: Wiley & Sons, Inc., 1959.

APPENDICES

APPENDIX A

DISTRIBUTION OF RAW SCORES ON THE OTIS QUICK-SCORING MENTAL ABILITY TEST (GAMMA TEST)

Raw Score	Frequency
72	3
71	0
70	2
69	1
68	2
67	1
66	2
65	4
64	1
63	3
62	2
61	5
60	2
59	2
58	7
57	3
56	4
55	1
54	3
53	5
52	6
51	7
50	4
49	5
48	3
47	0
46	4
45	0
44	4
43	0
42	0
41	1
40	1
39	1
38	0
37	1
36	1
35	1

APPENDIX B

EXPERIMENTAL GROUPS' INSTRUCTIONS

Instructions to Subjects Receiving Teaching Machine Presentation, One-Hundred Per cent Knowledge of Results and Self-Pacing

Before you is a "teaching machine". In the machine is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin make a full stroke upward with the handle on the left side of the machine (right side if you are left-handed). When this is done the first question will appear in the window on the top surface of the machine. You are to write your response on the answer sheet after reading the question. When you have made your response advance the machine by pushing the handle upward again. When this is done the correct answer and the next item will appear in the window. Compare your response with the correct answer and then read the next item and respond to it. Continue in this manner, at your own speed, until you are told to stop.

If, at any time, your machine does not function properly or you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response,

Appendix B (Contd.)

whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Teaching Machine Presentation, Partial Knowledge of Results and Self-Pacing

Before you is a "teaching machine". In the machine is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin make a full stroke upward with the handle on the left side of the machine (right side if you are left-handed). When this is done the first question will appear in the window on the top surface of the machine. You are to write your response on the answer sheet after reading the question. When you have made your response advance the machine by pushing the handle upward again. When this is done the next item will appear in the window and the correct answer to the previous item may or may not appear. If the correct answer to the previous question does appear, compare your response to the item with the correct answer. Then read the next item and respond

Appendix B (Contd.)

to it and move to the next item. Continue in this manner, at your own speed, until you are told to stop.

If, at any time, your machine does not function properly or you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Teaching Machine Presentation, One-Hundred Per Cent Knowledge of Results and Experimenter-Pacing

Before you is a "teaching machine". In the machine is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin make a full stroke upward with

Appendix B (Contd.)

the handle on the left side of the machine (right side if you are left-handed). When this is done the first question will appear in the window on the top surface of the machine. You are to write your response on the answer sheet after reading the question. When you have made your response wait until you are told to move to the next item, then advance the machine by pushing the handle upward again. When this is done the correct answer and the next item will appear in the window. Compare your response with the correct answer and then read the next item and respond to it. Each time you are instructed to go to the next item you should advance the next question by pushing the handle on the machine upward. Move on to the next item when you are told even if you have not finished the item you are working on. If you have time left over in the interval after responding to the item, continue to look at the item. Wait until you are told to advance before moving to the next item. The purpose of this is to ensure that each subject has equal exposure time to each item.

If, at any time, your machine does not function properly or you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

Appendix B (Contd.)

2. The abbriviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Teaching Machine Presentation, Partial Knowledge of Results and Experimenter-Pacing

Before you is a "teaching machine". In the machine is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin make a full stroke upward with the handle on the left side of the machine (right side if you are left-handed). When this is done the first question will appear in the window on the top surface of the machine. You are to write your response on the answer sheet after reading the question. When you have made your response wait until you are told to move to the next item, then advance the machine by pushing the handle upward again. When this is done the next item will appear in the window and the correct answer to the previous itme may or may not appear. If the correct answer to the previous question does appear, compare your response to the item with the correct answer. Then read the next item and respond to it. Each time you are instructed to go to the next item you should advance the next question by pushing the handle of the machine upward. If you have time left over

Appendix B (Contd.)

in the interval after responding to the item, continue to look at the item. Wait until you are told to advance before moving to the next item. The purpose of this is to insure that each subject has equal exposure time to each item.

If, at any time, your machine does not function properly or you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe thefollowing conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Programmed Text Presentation, One-Hundred Per Cent Knowledge of Results and Self-Pacing

Before you is a programmed textbook. In this book is a series of questions which you are to answer by responding on the answer sheet

Appendix B (Contd.)

provided. When you are told to begin turn to Item 1 at the top of page 2. The first question (Item 1) appears at the top of the page. You are to write your response to the question on the asnwer sheet. When you have done this turn the page; the correct answer to the previous question is at the top left of the page and the next item you are to respond to is directly to its right. Compare your response to the previous question with the correct response and then respond to the next item. Again you will find the answer on the next page along with the next item. Continue this process until you are directed back to page 1 and then go through the items directly below those you have completed. Continue in this manner until you have completed all seven rows of items. After completion of all seven rows of items go to Item 1 of Set 2 on page 10 and proceed through Set 2 in the same manner. Continue this process, at your own speed, until you are told to stop.

If, at any time, you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

Appendix B (Contd.)

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Programmed Text Presentation, Partial Knowledge of Results and Self-Pacing

Before you is a programmed textbook. In this book is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin turn to Item 1 at the top of page 2. The first question (Item 1) appears at the top of the page. You are to write your response to the question on the answer sheet. When you have done this turn the page. At the top of the next page is the next item you are to respond to; the correct answer to the previous question may or may not be at the immediate left of this question. If the correct answer to the previous item is present, compare your response with it before responding to the next item. Again you will find the next item at the top of the next page while the answer to the preceding item may or may not be present. Continue this process until you are directed back to page 1 and then go through the items directly below those you have completed (second row). Continue in this manner until you have completed all seven rows of items. After completion of all seven rows of items go to Item 1 of Set 2 on page 10 and proceed through Set 2 in the same manner. Continue this process, at your own speed, until you are told

Appendix B (Contd.)

to stop.

If, at any time, you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Programmed Text Presentation, One-Hundred Per Cent Knowledge of Results and Experimenter-Pacing

Before you is a programmed textbook. In this book is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin turn to Item 1 at the top of page 2. The first question (Item 1) appears at the top of the page. You are to write your response to the question on the answer sheet. When you have made your response wait until you are told to move to the next item,

Appendix B (Contd.)

then turn the page. The correct answer to the previous question is at the top left of the page and the next item you are to respond to is directly to its right. Compare your response to the previous item with the correct response and then respond to the next item. Each time you are instructed to go to the next item, turn the page, compare your answer to the correct answer at the top lefthand side of the page and respond to the next question. Continue in this manner until you are directed back to page 1, and then go through the items directly below those you have completed (second row). Upon completion of all seven rows of items you will be directed to Item 1 of Set 2 on page 10. Then proceed through Set 2 in the same manner.

If you have time left over in the interval after responding to an item, continue to look at the item. Wait until you are told to move to the next item before turning to the next page. The purpose of this is to insure that each subject has equal exposure time to each item.

If, at any time, you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "_____" indicates a one-word response, whereas "_____ _____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

Appendix B (Contd.)

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

Instructions to Subjects Receiving Programmed Text Presentation, Partial Knowledge of Results and Experimenter-Pacing

Before you is a programmed textbook. In this book is a series of questions which you are to answer by responding on the answer sheet provided. When you are told to begin turn to Item 1 at the top of page 2. The first question (Item 1) appears at the top of the page. You are to write your response to the question on the answer sheet. When you have made your response wait until you are told to move to the next item, then turn the page. At the top of the next page is the next item you are to respond to; the correct answer to the previous item may or may not be at the immediate left of this question. If the correct answer to the previous item is present, compare your response with it before responding to the next item. Each time you are instructed to go to the next item turn the page, compare your response to the correct answer, if it is present, at the top left hand side of the page and respond to the next question. Continue in this manner until you are directed back to page 1 and then go through the items directly below those you have completed (second row). Upon completion of all seven rows of items you

Appendix B (Contd.)

will be directed to Item 1 of Set 2 on page 10. Then proceed through Set 2 in the same manner.

If you have time left over in the interval after responding to an item, continue to look at the item. Wait until you are told to move to the next item before turning to the next page. The purpose of this is to insure that each subject has equal exposure time to each item.

If, at any time, you are uncertain as to what you are to do please notify the proctor immediately and ask for assistance.

Observe the following conventions:

1. The number of words needed to complete an item is indicated by the number of blanks. Thus "____" indicates a one-word response, whereas "____ ____" indicates a two-word response. When asterisks (***) are used in place of blanks, fill in as many words as you think necessary to respond to the item.

2. The abbreviation TT calls for a technical term. When it is used, a nontechnical word is incorrect.

3. There are often several reasonably equivalent responses, and it would be a waste of time to list them all. This is particularly true when the response is nontechnical. Use reasonable judgment in deciding whether your response is synonymous with the printed form. Score it correct if it is.

APPENDIX C

CRITERION EXAMINATION

1. The response system involved in walking to the door and opening it is _____ muscles.

2. In respondent behavior the response is _____ while in operant behavior the response is _____.

3. The emotional state of _____ is marked by such reflexes as sweating, contraction of blood vessels, "goose flesh", etc.

4. A generalized reinforcer is nearly _____ of the specific state of deprivation of the organism.

5. In a reflex, the _____ of a stimulus is the intensity which is barely sufficient to elicit a response.

6. A psychologist fed a baby when he emitted "coos" but not when he cried. We would expect that crying when hungry would be _____ because of the withholding of food.

7. Other things being equal, an operant which has been observed to occur at a high rate in the past has a high _____ of occurring at some future time.

8. Salivation elicited by the words "ice cream" is a(n) _____ response.

9. The initials GSR stand for _____ _____ _____.

10. In _____ behavior a stimulus precedes the response.

11. To get an animal to emit a response more frequently, we _____ the response.

12. When a previously neutral stimulus acquires the power to elicit a response it is called a(n) _____ stimulus.

13. The response system involved in passing food into the stomach is _____ muscles.

Appendix C (Contd.)

14. The rate of an operant can be increased by ending a(n) _____ _____.

15. In conditioning a reflex, as the number of pairings of conditioned and unconditioned stimuli increases, the _____ of the conditioned reflex decreases until it reaches a limit.

16. Establishing a conditioned reinforcer is similar to respondent conditioning in that both require that two simuli be _____.

17. In a reflex, the more intense the stimulus, the shorter the _____ of the reflex will be.

18. Turning on a very funny television program is reinforced by the presentation of a(n) _____ reinforcer.

19. A special instrument, called a _____, is used to measure the electrical resistance of the skin.

20. In the extinction process of the conditioned reflex, the _____ _____ is presented alone.

Test B

1. Two ways of effectively avoiding undesirable conditioned responses are to extinguish them or to condition a(n) _____ response to the same situation.

2. The response system which provides the stomach with digestive juices is the _____.

3. A reinforcement which consists of terminating a painful stimulus is called _____ reinforcement.

4. The process by which a conditioned stimulus loses its power to elicit the conditioned response is called _____.

5. To condition a reflex, a neutral stimulus is paired with a(n) _____ _____.

6. In conditioning a reflex, as the number of pairing of conditioned and unconditioned stimuli increases, the _____ of the conditioned response increases until it reaches a limit.

7. In "detecting a lie", a "lie detector" measures responses which became conditioned through the pairing of lying and _____.

Appendix C (Contd.)

8. When a response is elicited by a stimulus without previous conditioning, the sequence is called a(n) _____ _____.

9. _____ behavior is influenced by the consequences of previous, similar responses and in this type of behavior a stimulus does not necessarily precede the response.

10. Smooth muscles change the dimensions of various _____ organs.

11. Any specific instance of an operant is called a(n) _____.

12. When a man's hand is touched by a hot surface or receives an electric shock, the hand is immediately withdrawn. The withdrawal of the hand is a(n) _____ _____.

13. In a reflex, the more intense the stimulus, the greater the _____ of the response will be.

14. _____ is due to a passage of time during which the response is not emitted.

15. A stimulus which acquires the property of a reinforcer is called a(n) _____ reinforcer.

16. The process of conditioning was discovered by a man named _____.

17. In psychology the technical term for "reward" is _____.

18. Food given to a hungry animal does not reinforce a particular response unless it is given immediately _____ the response.

19. If a chimpanzee exchanges tokens for food, water, a mate, and escape from pain, a token becomes a(n) _____ _____.

20. Certain groups of responses, such as those elicited by a sudden loud noise, are characteristic of a state of _____.

APPENDIX D

RAW DATA

Subject No.	Trt	PT		IC	E	ER	TER
1	1002	16 14	30	334	7	2	8.1
2	1002	16 15	31	334	11	3	10.0
3	1001	4 9	13	334	95	28	31.9
4	1001	8 9	17	334	124	37	37.5
5	1000	7 7	14	334	116	35	36.3
6	1000	8 10	18	334	124	37	37.5
7	1012	10 10	20	253	16	6	14.2
8	1012	7 7	14	222	23	10	18.4
9	1011	4 5	9	155	29	19	25.8
10	1011	7 6	13	221	38	17	24.4
11	1010	9 5	14	218	61	28	31.9
12	1010	3 6	9	152	41	28	31.9
13	1102	6 9	15	334	105	31	33.8
14	1102	15 12	27	334	60	18	25.1
15	1101	14 11	25	334	36	11	19.4
16	1101	10 9	19	334	49	15	22.8
17	1100	6 7	13	334	88	26	30.7
18	1100	9 12	21	334	93	28	31.9
19	1112	13 13	26	379	7	2	8.1

Appendix D (Contd.)

Subject No.	Trt	PT		IC	E	ER	TER
20	1112	9 12 15	21	379	7	2	8.1
21	1111	15 10	30	345	20	6	14.2
22	1111	8 5	18	254	35	14	22.0
23	1110	6 6	11	200	42	21	27.3
24	1110	8 9	14	209	74	35	36.3
25	0002	8 14	17	334	144	43	41.0
26	0002	12 7	26	334	51	15	22.8
27	0001	6 12	13	334	191	57	49.0
28	0001	14 12	26	334	70	21	27.3
29	0000	10 4	22	334	107	32	33.8
30	0000	6 11	10	334	150	45	42.1
31	0012	9 13	20	264	11	4	11.5
32	0012	15 5	28	329	20	6	14.2
33	0011	5 8	10	240	36	15	22.8
34	0011	9 12	17	226	40	18	25.1
35	0010	8 7	20	241	15	6	14.2
36	0010	8 12	15	159	24	15	22.8
37	0102	11 9	23	334	45	13	21.1
38	0102	10 9	19	334	119	36	36.9
39	0101	6 13	15	334	81	24	29.3
40	0101	12 10	25	334	33	10	18.4
41	0100	8 7	18	334	157	47	43.3
42	0100	7	14	334	167	50	45.0

Appendix D (Contd.)

Subject No.	Trt	PT		IC	E	ER	TER
		16					
43	0112	14	30	316	33	10	18.4
		11					
44	0112	10	21	279	0	0	0
		6					
45	0111	7	13	214	71	33	35.1
		11					
46	0111	9	20	363	17	5	12.9
		7					
47	0110	7	14	212	40	19	25.8
		6					
48	0110	3	9	230	100	43	41.0

Headings are:

Trt = treatment combination*

PT = criterion test score

IC = items completed

E = errors on program

ER = error rate

TER = transformed error rate

*First digit refers to mode of presentation
0 = programmed text presentation
1 = teaching machine presentation

Second digit refers to schedule of reinforcement
0 = partial reinforcement
1 = 100 per cent reinforcement

Third digit refers to type of pacing
0 = experimenter-paced
1 = self-paced

Fourth digit refers to intellectual level of subject
0 = lowest intellectual level
1 = intermediate intellectual level
2 = highest intellectual level

VITA

Gary W. Evans

Candidate for the Degree of

Doctor of Philosophy

Thesis: MODE OF PRESENTATION, PACING, KNOWLEDGE OF RESULTS AND INTELLECTUAL LEVEL IN AUTOMATED INSTRUCTION

Major Field: Psychology

Biographical:

Personal Data: Place of birth - Beaver, Oklahoma; Date of Birth - November 18, 1935.

Education: Bachelor of Science degree, June, 1957, Kansas State College at Pittsburg, major - Psychology; Master of Science degree, June, 1959, Kansas State College at Pittsburg, major - Psychology; Doctor of Philosophy degree, requirements completed, May, 1963, Oklahoma State University, major - psychology.

Experience: Graduate Teaching Assistant, September, 1959-May, 1962, Oklahoma State University, Stillwater, Oklahoma; Research Intern, June, 1960-August, 1960, Human Resources Research Office, Fort Bliss, Texas; Assistant Professor of Psychology, September, 1963-present, Fort Hays State College, Hays, Kansas.