DOCUMENT RESUME

ED 113 672 CG 010 177

AUTHOR Parker, Reese
TITLE The Effects of Amount of Study Time, Quality of Instruction, and Degree of Initial Learning Upon Retention of an Intellectual Skill (Summary).
PUB DATE [75]
NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975); Not available in hard copy due to marginal reproducibility of original document

EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Achievement; *Instructional Programs; Intelligence; *Learning; Post Testing; Research Projects; *Retention Studies; *Skill Development; *Study Skills

ABSTRACT

The present study sought to discover the effects on delayed retention of two different degrees of quality of initial instruction for two different amounts of learning time; which combination of quality and time is most likely to be achieved in normal school situations; and the implications for those who design instructional materials. Two programed texts, one superior to the other in the quality of produced level of initial learning, were prepared and administered to 120 high school seniors randomly assigned to one of four equal-sized treatment groups, in combination with each of two differing amounts of learning time. An immediate posttest and a delayed retention test were administered. Parametric and nonparametric statistical tests yielded the following results: (1) an increased amount of study time promoted differing levels of achievement in delayed retention; (2) existing differences among groups could be attributed to effects of differing quality of instructional levels and/or amount of initial learning; and (3) no interaction effects were significant. Limitations of the study are delineated. (Author/SE)

ED113672

THE EFFECTS OF AMOUNT OF STUDY TIME, QUALITY OF INSTRUCTION, AND DEGREE OF INITIAL LEARNING UPON RETENTION OF AN INTELLECTUAL SKILL

by
REESE PARKER

WEBER STATE COLLEGE

(SUMMARY)

AERA ANNUAL MEETING
Washington, D.C.
1975

26,04

CG 010 172

INTRODUCTION

Overview

Sizeable amounts of resources are necessary to maintain an extensive public education system. Support for these expenditures is forthcoming, in part, because the system assumes long term retention of skills as a goal. Such support would be less likely if only short term retention resulted from the system, as this would preclude long term benefits for the society which supports education. Therefore, an important part of the justification for the extensive educational endeavor rests upon the system's capability for promoting recall and use of skills, concepts, and principles, by students, over extended periods of time.

How would an educational system promote this capability for long-term recall of essential skills in its instructional efforts? While various possibilities for accomplishing this exist, three are at once apparent: a) provide for instruction which promotes highly effective initial learning on the part of students served by the system, b) allow more time for students to engage in the learning of these essential skills, and c) provide for strategically placed review and rehearsal periods following initial instruction in essential skills. Of these alternatives, two (a and b) were selected for investigation in this study.

The study reported here was concerned with the investigation of retention of intellectual skills. The experimental treatments provided instruction which promoted two different amounts of learning time. Formative evaluation techniques were used to develop the instruction to provide differing levels of initial learning. The content of the experimental instruction was concerned with developing skills in students which would enable them to apply a method, developed by Toulmin (1958), in analyzing positions taken on social issues. In terms of goals of instruction in social studies, skills needed to analyze positions taken on social issues are assigned high priority by Allen and Adair (1971), Fenton (1967), and Massialas and Smith (1965), as well as other authorities in the field.

Programmed texts were used as the delivery system for the experimental instruction to avoid uncontrolled teacher-related influences.

Problem Statement

The problem investigated here concerned two issues which are directly relevant to planning, designing, developing, and implementing instruction which will insure retention. The two issues of concern

are closely interrelated. As was stated earlier, two immediately apparent possibilities for promoting retention of instruction are to a) increase the effectiveness of initial learning by improving the quality of instruction and b) allow students more time for whatever instruction is offered.

In implementation of instruction, insuring retention would be more likely if students increased initial learning, i.e. if they had access to instruction which promoted effectiveness in initial learning. Also, if students were allowed to participate in the instruction for periods of time sufficient to increase initial learning, it would seem that retention would be likely to increase. In light of these two hypotheses, it would seem appropriate, in the implementation of instruction, to employ instruction that insured a high level of initial learning effectiveness and/or to provide sufficient time for the student to engage in instruction to increase effectiveness of initial learning. The question remains as to which of these two avenues of improvement is most practical, and how the two might interact.

So the second issue is, "How do we improve instruction to fulfill needs implied by the two postulates?" To address this question, several others must first be faced. First, should instruction have associated with it, prior to implementation, a known level of effectiveness for insuring initial learning? Second, in the realities of most instructional environments, is it feasible to determine effectiveness of instruction prior to implementation? Third, if this latter practice is feasible, how effective should instruction be before it is used? Finally, if instruction proves to be of a lower level of effectiveness than is desired, how can it be feasibly improved, or, how can instructional time be utilized to compensate for a lower level of effectiveness?

The problem then, revolves around the possibilities of increasing retention levels of students as a result of manipulation of a) the time students are engaged in instruction, and/or b) the level of effectiveness regarding initial learning associated with the instruction students engage in. Specifically, the investigation described here sought to examine the following experimental questions:

> What are the effects upon delayed retention of two different degrees of quality of initial instruction for two different amounts of learning time? Which combination of quality and time is likely to be most practical to achieve in normal school situations? What are the implications for those who design instructional materials?

RELATED STUDIES

The results of other studies justify the choice of time allowed for instruction and quality of instruction from the Carroll model as variables in the present study of retention of purposeful instruction. This is also true of the variable, degree of initial learning, even though we do not define this as Carroll does.

Quality of Instruction

It is difficult to locate studies where comprehensive revision of instruction for improved quality is considered as an experimental variable. According to Sulzen (1972), when experimenters report the use of differing quality levels of instruction in their studies, they most often adjust only one of many possible components of instruction, e.g., clarify directions, give more examples, use more illustrations, give objectives to one group and withhold objectives from another, etc. The lack of studies dealing with entire composites of such components serves to emphasize the need for such experiments as this one. While it is profitable to study such single components as those just cited, this does not replace the need for varying total effectiveness of composites.

Time Allowed for Instruction

Several studies have investigated time allowed for instruction as an independent variable. An experiment by Deady (1969) was concerned with the effects of increased time allotments on students' achievements and attitudes in science classes. Students in experimental groups spent 35 days in a unit; control groups were allowed only 20 days. The use of teacher-dependent instruction with several teachers removed the possibility of controlling for teacher based variance. An analysis of results from common tests revealed no significant differences between control and experimental groups regarding either attitude or achievement.

Corrozi (1970) investigated effects of increased reading time, different types of questions, and differing instructional formats on short-term and delayed retention. He found that reading time had a positive effect upon both short-term and delayed retention of meaningful prose materials, but not for nonsense materials. This agrees with Ausubel's (1970) conclusion that presentation in a meaningful context is necessary for the delayed retention of instruction, but context exerts no positive effect in retention of nonsense material.

A *post hoc* study by Jarvis (1965) found significant superiority for increased time allotments on immediate tests in reading, arithmetic, and language arts achievement, but not on tests of delayed retention. This study involved 713 sixth graders in 35 schools, so the results may have been heavily influenced by lack of experimental control and by teacher-based variance.

Wang and Lindvall (1969), in performing a validation study of the Carroll model, found little relationship between aptitude and learning rate (time taken in study to pass a mathematics skills test) and no significant differences between groups taking short and lengthy periods of time in arriving at acceptable performance levels on the test.

In a study involving 2,000 twelfth graders using Harvard Project Physics materials, Welch and Bridgham (1968) reported no significant differences, regarding retention, between groups who studied the materials for differing time intervals. Teachers selected intervals used, types of instructional activities, and ratios of types of instructional activities students were involved in. Therefore, very little experimental control was exerted in the study.

Carroll (1973) used aptitude and achievement data gathered from standardized testing in a nationwide study to develop a model such that, a) for students in any grade, expected achievement is directly proportional to measured aptitude, and b) the proportionality coefficients (achievement/aptitude) are a function of time, except for some cases involving first and second graders. This model demonstrates that achievement at time _t_, the achievement of the average individual, is the average achievement for all individuals at that time. Therefore, since the predictor of achievement from aptitude is time-correlated, individuals with slow learning rates (who need more time to learn) will have less than average achievement, and those with rapid learning rates will have greater than average achievement scores.

To summarize, then, among students concerned with effects of increased time allotments, it can be noted that results are mixed. While these mixed results often seem to result from lack of standardization of conditions within experiments, it could also be that the increased time intervals employed were not sufficiently extended to account for variance that occurs as a result of including slower learners in the experiments.

Degree of Initial Learning

In a review of research regarding retention, Seagoe (1970) proposed to teachers that the single most important factor in preventing forgetting is degree of initial learning. She suggested that teachers enhance retention by using a) techniques which promote over-learning, b) distributed practice during instruction, and c) application of initial learning in differing situations following instruction. This emphasis upon attempting to promote the highest degree of initial learning in order to promote retention is in concurrence with guidelines suggested by Gagne and Rohwer (1969) and Gagne (1970). Similar suggested lists of retention-promoting practices designed to increase degree of initial learning have been offered by Davis (1966), who adds mental practice (learner rehearsal) to his list, and Krumboltz (1964) who suggests the superiority of overt responding over covert responding.

In a study of effects of three types of repetition upon retention, Winne et al (1972) found the amount of initial learning to be a better predictor of retention than any type of repetition involved in the experiment, for both specific fact retention and "chunk" retention. This result held for information presented in meaningful context and lists of facts presented in random arrangement.

Bobrow and Bower (1969) found that initial learning and retention could be increased if students constructed sentences with criterion material to provide a personalized basis for recall. This use of composing of sentences as a method for a construction strategy, which the writers identified as a mathemegenics, is not of great import to the present study, but the great differences in degree of original learning resulting from their use is. This difference was reflected in superiority in retention for groups with greater initial learning.

Several studies have produced results contrary to the findings just cited regarding the importance of initial learning in promoting retention. Reynolds and Glaser (1964) found that there was no relationship between the amount of original learning and degree of retention, when the differences in original learning resulted from differing numbers of examples used in two versions of a programmed text in mathematics. Gibson (1969) found that over-learning and drill with rules promoted initial learning, but exerted no significant effect upon delayed retention. Gay (1971), using time for relearning as the measure of retention, found no savings in the number of examples necessary in the original learning to reach an arbitrary criterion in problem solving using mathematics rules.

In summary, while the weight of evidence seems to substantiate the importance of initial learning for promoting delayed retention, these last three studies resulting in contrary findings dealt exclusively with intellectual skills and might be more important in interpreting results of the study described here, as they tend to support Gagne's feeling (personal communique) that students either have mastery of intellectual skills or they do not. That is, mastery for discrete intellectual skills is likely to be absolute; it is present in the learner or it is not. On the other hand, overlearning of information may enhance retention.

Implications for This Study

The implications of the preceding interpretations hold importance for the design of instructional materials and the assessment of delayed retention in this study due to the following factors. First, if the conditions which have been identified as promoting retention can be designed into the experimental instruction to be used, the possibility of designer bias in favor of any treatment will be less likely to occur. Second, if the effect of degree of initial learning upon delayed retention can be validly accounted for by adjusting retention scores, either by using the formula suggested by Davis (1966),

or statistically as suggested by Wodtke (1967) and Cook (1970), the effects of amount of time allowed for learning can be more clearly isolated, as can the effects of quality of instruction.

The mixed effects upon retention reported here for increased time allotments and degree of initial learning, coupled with the seeming absence of studies employing a holistic definition of quality instruction, especially when quality is not affected by teacher-dependent activity, serves to emphasize the need for studies such as the one described here.

METHOD

Overview

The developmental and experimental procedures involved in implementing this study were performed as follows. First, a programed text (referred to in later sections as version I) dealing with the use of Toulmin's (1958) schema for analyzing position statements was developed and implemented with a group of 30 high school seniors. Second, data gathered from student comments, test results, and experimenter analysis of these data, were used to revise the version I text yielding a second text (referred to in later sections as version II). Third, version II was implemented with a second group of 30 high school seniors and data were gathered to insure that version II was superior to version I in the level of initial learning that version II produced for students.

Fourth, 120 high school seniors from two large English classes were randomly assigned to one of four equal-sized treatment groups. Fifth, each of the two versions of the programed text was administered to one of the four treatment groups, in combination with each of two different amounts of learning time. Sixth, when alloted study time expired for each group, an immediate posttest was administered to its members. Due to length of time required to complete the fifth and sixth procedures, the 120 students involved reported to a large lecture classroom for 90 minutes on each of two consecutive days. The first day was required for completing the programmed text and the second was used for completing the programmed text for the larger study time and for testing. Seventh, a delayed retention test was administered to the 120 students in a single group, 13 days after the immediate posttest. Finally, data obtained from the two tests were subjected to parametric (analysis of variance) and nonparametric (Wilson Chi-Square) analyses to test the effects of a) quality of instruction level, b) time engaged in instruction, c) level of initial learning, and d) interactions among a, b, and c, upon delayed retention of skills involved in applying the Toulmin schema.

Experimental Design

A 2 x 2 design with repeated measures was employed in the study using two levels of quality of instruction and two amounts of study time for each quality of instruction. The two levels of quality of instruction were achieved and documented by implementing two formative evaluations. The first of these documented the quality of the first draft instructional materials and provided data for revising them. The second allowed documentation of the quality level of the

revised instruction. Two versions of a programmed text were employed to deliver the instruction, each version containing one of the levels of quality which resulted from formative evaluation.

An immediate posttest and a delayed retention test comprised the repeated measures factor in the design as each was a measure of retention evident at a different point in time. In applying parametric analysis of variance techniques to data obtained from the experiment, the immediate posttest (degree of initial learning) was treated as an independent variable as suggested by Wodtke (1967). Such a practice is appropriate, according to that writer, as measures of initial learning and delayed retention measures are usually highly correlated. When a measure of delayed retention is used as the dependent variable, as was the case in the present study, consideration of the immediate performance test (degree of initial learning) as an independent variable enables the experimenter to: a) determine the degree of the relationship between the two repeated measures, b) account for effects of individual differences between participants in the experiment, and thereby, c) obtain a more accurate estimate of the overall effects of treatment variables in the experiment.

Several researchers have determined that a repeated measures design is an appropriate one for analysis of data in retention studies. Campbell and Stanley (1963), Wodtke (1967), and Cook (1970) have given attention to the potential of repeated measures designs, and each of these authors cites four strengths inherent in such designs for retention studies.

First, repeated measures designs provide control for confounding that could be exerted by carry-over effects. Such effects often result from the order in which observations are obtained in, say, factorial designs.

Second, in retention experiments, differences among individuals are often quite large relative to differences the experimenter is attempting to evaluate. The repeated measures design provides statistical controls for individual differences; these controls ameliorate difficulties concerned with relative sizes of differences existing in an experiment.

Third, if an additive model is realistic in an experiment, it is likely that sequence effects will interfere with the determination of differences. Such interference is also controlled in the repeated measures design.

A fourth strength cited for repeated measures designs results from the cumulative effects of the three previously cited strengths. Due to this cumulative effect, the degree of precision of parameter estimates increases more rapidly than in other designs as the numbers of subjects per treatment increases when repeated measures are used.

These four strengths are not characteristic of two other designs which are often employed in retention experiments: counterbalanced designs and analysis of covariance. Thus, a repeated measures design would seem to be most appropriate for analysis of data resulting from implementation of this study.

Data obtained from the experiment were subjected to both parametric (analysis of variance) and nonparametric (Wilson Chi-Square) analyses because: a) there is continuing controversy regarding proper conditions for their respective uses, b) it is likely that assumptions required for parametric techniques were violated in this study, and c) parametric techniques are usually superior to nonparametrics in power comparisons for the same sample size when the parametric assumptions are met.

Analysis of Data

Parametric and Nonparametric Components

Data obtained from the study were analyzed using both parametric and nonparametric methods. The rationale for this procedure, presented in a previous section, focused upon three issues; (a) the continuing controversy regarding proper conditions for use of parametric and nonparametric techniques, (b) the likelihood, in strict interpretation, that conditions said to be necessary for using parametric techniques were violated in implementing this study, and (c) the fact that parametric techniques are usually favored over nonparametric techniques in power comparisons for the same sample size when parametric assumptions are met.

A three-factor model was employed in the parametric analysis, with repeated measures on the last factor. It can be stated as follows:

$$Y_{ijkl} = M + T_i + P_j + A_k + TP_{ij} + TA_{ik} + PA_{jk} + TPA_{ijk} + R_{l(ij)} + RA_{kl(ij)}$$

Where:

T = time engaged in instruction; 2 levels; fixed.
P = quality of instruction; 2 levels; fixed.
A = performance assessment; 2 levels; immediate and delayed.
R = replications; 30 Ss per treatment; random.

The two levels of time engaged in instruction were 75 and 105 minutes. Quality of instruction levels corresponded to immediate performance levels obtained using each of two versions of the instructional materials, and the two levels of performance assessment employed were immediate and delayed.

The nonparametric technique used was a distribution-free test for analysis of variance hypotheses developed by Wilson (1956). This technique was derived from an earlier work by Rao (1952) which demonstrated that a Chi-square statistic for a contingency table can be broken down in much the same way as a total sum of squares is decomposed in parametric analysis of variance computations. It is applicable to problems involving designs of two or more variables. The general formula used in this procedure is:

$$X_T^2 = \sum_{ijk}\sum\sum \left[\frac{(afijk - n_{ijk}\, n_a/n)^2}{n_{ijk}\, n_b/n} + \frac{(bfijk - n_{ijk}\, n_b/n)^2}{n_{ijk}\, n_b/n} \right]$$

Where:

n_a = number of observations greater than or equal to the overall median on the delayed performance assessment.

n_b = number of observations less than the median on the delayed performance assessment.

afijk = number of observations in a cell greater than or equal to the overall median on the delayed performance assessment.

bfijk = number of observations in a cell less than the overall median on the delayed performance assessment.

n_{ijk} = total observations in a cell, both above and equal to, and below the overall median on the delayed performance assessment.

This technique is not adaptable to repeated measures design, so the analagous parametric statistical model for the nonparametric analysis is stated in the form of a three factor factorial design, as follows:

$$Y_{ijkl} = M + T_i + P_j + A_k + TP_{ij} + TA_{ik} + PA_{jk} + TPA_{ijk} + R_{l(ijk)}$$

Methods of Data Analysis

Data resulting from the study were subjected to an analysis of variance for the parametric analysis using computer program 08V from the UCLA biomedical series (Dixon, 1970). Output from the program includes cell means, cell deviations, and an analysis of variance source table which includes: sources of variance, sums of squares, degrees of freedom, mean squares, and expected mean squares for terms specified in the model. The parametric ANOVA output for this study is presented in Table 1 and descriptive data for the experiment is presented in Table 2.

In performing F - tests for the parametric analysis, an alpha level of .05 was deemed acceptable with regard to the probability of committing a Type I error in decisions concerning rejection of the null hypotheses.

For the nonparametric technique, the data were cast in a 2x2x2x2 (2 x r x c x b) contingency table. Rows in the contingency tables represented the two levels of quality of instruction in the study. Columns represented the two levels of time engaged in instruction used in the study. Blocks represented the effects of initial learning indicated by individuals' scores on the immediate performance assessment.

Subjects were assigned to cells in the contingency table according to the following procedure. The overall medians for distributions of scores on the immediate and delayed performance assessments, respectively, were determined. Subjects with scores equal to or above the median on the delayed performance assessment were assigned to one portion of the contingency table (n_a), and those with scores on the same measure which were below the overall median were assigned to another portion of the table (n_b). Depending upon the treatment group they had been assigned to, they were subsequently assigned to a cell in the appropriate portion of the table depending upon whether their scores on the immediate performance assessment were above or equal to, or below, the overall median on that measure. The results of implementing these cell assignment procedures for this study are presented in Table 3.

In performing Chi-square tests for the nonparametric analysis, an alpha level of .05 was deemed acceptable with regard to the probability of committing a Type I error in decisions concerning rejection of the null hypotheses.

Table 4 presents results of the Chi-square tests which were performed in the course of the nonparametric analysis. Obtained Chi-square values are not presented for discrete interaction effects because the total possible value of Chi-square attributable to interaction was not statistically significant for the two-way interaction with the least degrees of freedom. Therefore, there could be no significant interaction effect in the nonparametric analysis.

The reader should be aware that while each of the two statistical analyses used in the study was appropriate, they were not parallel. The Wilson Chi-Square did not allow use of repeated measures, therefore the results of the tests should be viewed as being derived from two separate analyses. While the results of implementing the two techniques will be discussed together in sections that follow and conclusions inferred from the two tests are discussed as being related, there is no assumption that the two analyses are actually the same.

RESULTS

Effects of Differing Time Levels in Instruction

The parametric null form of the statistical hypothesis tested for this experimental hypothesis was :

$$Ho: \quad \sigma^2_{Time} = 0.$$

Results of the analysis of variance were used to perform an F - test at alpha = .05, with 1 and 116 degrees of freedom. The rejection region for this test is represented by obtained F values greater than 3.93. The obtained F for this hypothesis was 6.8, therefore, the null hypothesis was rejected. The effect of differing time levels in the experiment was found to be significant. According to information provided by Cohen (1969), power for this statistical test was found to be approximately .97 for a moderate effect size.

The nonparametric null form of the statistical hypothesis tested was: Ho: There are no differences among groups which can be attributed to effects by differing time levels in instruction.

Results obtained were tested using a Chi-square statistic with alpha = .05, and 1 degree of freedom. The rejection region for this test consisted of Chi-square values greater than 3.84. The obtained Chi-square value for this test was 1.6 (power = .93 for a moderate effect size) which did not allow rejection of the null hypothesis. Therefore, the effects of using differing time levels in the experiment was found to be not significant in the nonparametric test.

The decision to reject the null hypothesis in the parametric case, meant that the alternative hypothesis ($Ha: \sigma^2\ Time \neq 0$) could not be rejected. This decision supports the contention that differing levels of achievement in delayed retention were promoted by an increased amount of study time. Failure to reject the null hypothesis in the nonparametric test disallowed any inference with respect to the alternative hypothesis.

These tests proved to be the only occasion in the experiment where similar statistical decisions were not supported by results of both parametric and nonparametric tests. In the nonparametric analysis of data, immediate posttest scores were arbitrarily divided at the median of that distribution (see Table 3) to provide for use of amount of initial learning as a variable in the analysis. This decision, while consistent with directions for use of this particular Chi-square technique

(Wilson, 1956), promoted manifestation of one issue of the continuing parametric/nonparametric controversy in the results of this experiment. That is, valid parametric tests are more sensitive to differences and therefore, are more powerful for the same experimental conditions than are similar nonparametric techniques, when all parametric assumptions are met.

In preparing the data for the nonparametric analysis, the level of measurement necessarily used was nominal as opposed to the interval level used in ANOVA. The particular classification system employed in the Wilson Chi-square test calls for dividing scores into two groups; those below the median and those above or equal to the median. Thus, differences in individual scores, on the order of, say, 10 points on an interval scale, are not evidenced in the nominal scale employed in the nonparametric analysis. The reader may determine, for himself, which of these data scale assumptions meets his criteria.

Effects of Differing Quality Levels in Instruction

The parametric null form of the statistical hypothesis tested was: $\sigma^2_{\text{Quality Levels}} = 0$. An F-test was performed at alpha = .05 and 1 and 116 degrees of freedom. The rejection region for this test consisted of obtained F values greater than 3.93. The obtained F for this test was 4.4, which allowed rejection of the null hypothesis, and the power level associated with the decision was approximately .97 for a moderate effect size.

The nonparametric statistical hypothesis tested was: Ho: There are no differences among groups which can be attributed to effects of differing quality of instruction levels. The obtained Chi-square value for this hypothesis was 4.4, which allowed rejection of the null hypothesis at a power level of approximately .93 for a moderate effect size. The rejection region was represented by obtained values greater than 3.84, for Chi-square with alpha = .05 and 1 degree of freedom.

Rej[illegible]ing the null hypothesis in the parametric and nonparametric tests meant that the respective alternative hypotheses (Ha: $\sigma^2_{\text{Quality Levels}} \neq 0$ and Ha: There are differences among groups which can be attributed to effects of differing quality of instruction levels) could not be rejected, i.e., the use of differing quality levels of instruction can promote achievement of differing delayed retention levels.

Effects of Amount of Initial Learning

The parametric statistical hypothesis tested was: $\sigma^2_{\text{Amount Initial Learning}} = 0$. The obtained F value for the hypothesis was 128.4. The rejection region consisted of obtained

F values at alpha = .05 with 1 and 116 degrees of freedom, greater than 3.83. The null hypothesis was rejected with a power level of approximately .97 for a moderate effect size.

The nonparametric statistical hypothesis tested (Ho: There are no differences among groups which can be attributed to effects of amount of initial learning) was rejected with an obtained Chi-square value of 25.3. The rejection region for the hypothesis consisted of obtained Chi-square values, with alpha = .05 and 1 degree of freedom, which were greater than 3.84. A power level of approximately .93 was associated with the decision to reject the null hypothesis for a moderate effect size.

Rejection of the null hypothesis in the parametric test meant that the alternative hypothesis $(Ha: \sigma^2_{\text{Amount Initial Learning}} \neq 0)$ could not be rejected. A similar conclusion with respect to the nonparametric null hypothesis was appropriate, therefore, the alternative hypotheses (Ha: There are differences among groups which can be attributed to effects of amount of initial learning) could not be rejected; i.e. that differences in amounts of initial learning promote differing levels of achievement in delayed retention.

Effects of Interactions Among Variables

Statistical null hypotheses for all interactions among the 3 independent variables were tested using parametric means. No obtained F was sufficiently large to allow rejection of any null hypothesis concerning interactions (see Table 3). The rejection region for F with alpha = .05 and 1 and 116 degrees of freedom consists of F values greater than 3.84. The largest F for any interaction hypothesis was 1.7. The power level associated with this statistical decision was approximately .92 for a moderate effect size.

As reported previously, no nonparametric test allowed rejection of any statistical null hypothesis concerned with interactions, as the obtained Chi-square value for total interaction (2.2) was not significant prior to partitioning it into the two- and three-way interactions. This negated the possibility that any discrete interaction could be found to be significant by nonparametric methods.

Limitations

There are three possible sources known that could exert limiting influences upon results in the present study. Two of the sources will be reported here and one will be reported later in the Unanticipated Outcomes section of this paper.

First, there is the possibility that confounding existed between the variables of quality levels of instruction, and the amount of

initial learning. Quality levels of instruction were obtained for the two versions of the programmed text by determining the proportion of subjects who achieved an arbitrary performance criterion level during the formative evaluation cycle. Therefore, when subjects were assigned to treatment groups, especially for the shorter time period (75 minutes), only a certain proportion of them (33% for version I and 63% for version II) were likely to achieve the arbitrary performance criterion level.

At the same time, amount of initial learning for each individual in the experiment was measured by the number of items correct on the immediate posttest. It is possible, assuming the logic of the Carroll model to be correct, that the combination of quality level of instruction and time allotment employed in each treatment group constituted an upper limit on the number of students who could achieve a degree of initial learning equal to the arbitrary minimum performance level. That is, for a given quality level of instruction and any time allotment that is less than what is needed for each individual in a group, it is only possible for some proportion of the group less than 100% to achieve a desired performance level. Confounding would result in that we may not have accurately determined relative effects, upon the immediate performance level achieved by an individual (as contrasted to the average student) of time allotment <u>versus</u> quality level of instruction.

The second possible source of limitations on the results of this study concerned a promotion being operated by a local radio station. Under the rules of the promotion, the school which turned in the most 3 x 5 inch cards with the words "Care Free Gum" written on them would win a concert at the school by an internationally known rock music group. The two days during which the programmed texts and immediate performance tests were implemented were also the final two days of the contest and subjects were quite interested in writing cards. While subjects did not engage in producing cards when they were working in the two versions of the programmed text, it was possible that they were not attending fully to the task.

Thirteen days later, when the experimenter returned at the appointed time to administer the retention test, the students were not in the large lecture classroom. An inquiry at the school office led to the discovery that the subjects were in the auditorium where they were being entertained by an internationally know rock group--a prize won by the school in the "Care Free Gum" contest. About twenty minutes later the subjects returned to take the retention test. It is possible that the "Care Free Gum" phenomenon exerted effects upon the results of this study to some unknown extent.

TABLE 1

Analysis of Variance

Source	d.f.	MS	F
Mean (M)	1	75650	------
Time engaged in instruction (T)	1	424	6.8*
Quality of instruction (P)	1	275	4.4*
Immediate performance assessment (A)	1	1799	128.4*
T X P	1	18	.3
T X A	1	4	.3
P X A	1	26	1.7
T X P X A	1	6	.4
Subjects (S) within treatments	116	62	------
S X A within treatments	116	15	------

* significant with alpha = .05

TABLE 2

Descriptive Data for Experiment

Group	Test	Range	Mean	Standard Deviation
version I	posttest	24	18.4	7.65
75 minutes	delayed retention	20	12.7	7.35
version II	posttest	26	20.8	5.58
75 minutes	delayed retention	14	15.9	4.60
version I	posttest	25	22.4	5.70
105 minutes	delayed retention	24	15.4	5.90
version II	posttest	26	26.4	6.85
105 minutes	delayed retention	22	17.8	6.65

TABLE 3

Nonparametric Contingency Table

		Time Engaged in Instruction			
		75 minutes		105 minutes	
		<Md immediate	>Md immediate	<Md immediate	>Md immediate
(above overall Md on delayed assessment) Quality of Instruction	Level 1 (Low)	3	6	3	12
	Level 2 (High)	7	9	1	16
(below overall Md on delayed assessment) Quality of Instruction	Level 1 (Low)	13	8	11	4
	Level 2 (High)	10	4	10	3

TABLE 4

Results of Nonparametric Analysis

Source	df	Chi-square
Total Chi-square	7	33.5*
Time engaged in instruction (T; column effect)	1	1.6
Quality of Instruction (P; row effect)	1	4.4*
Immediate performance assessment (A; block effect)	1	25.3*
Interaction effects (total)		2.2

* significant with alpha = .05

REFERENCES

Adair, D. H. & Allen, R. F. Violence and riots in urban America. New York: Charles Jones, 1969.

Adams, J. A. Human memory. New York: McGraw-Hill, 1967.

Allen, R. F. Learning about religion in American history courses. Tallahassee: Religion-Social Studies Curriculum Project, 1972.

Atkinson, R. C. & Shiffren, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Ed.), The psychology of learning and motivation. New York: Academic Press, 1968.

Bloom, B. S. Mastery learning. Chicago: University of Chicago, 1969. (mimeograph)

Bobrow, S. A. & Bower, G. H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Briggs, L. J. Handbook of procedures for the design of instruction. Pittsburgh: American Institutes for Research, 1970.

Broadbent. D. E. Perception and communication. New York: Pergammon, 1958.

Brown, J. Some tests of the decay theory of immediate memory. Quarterly Journal of Experimental Psychology, 1958, 10, 12-21.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, Spring, 1970, 71-80.

Carroll, J. B. Importance of the time factor in learning. Paper presented at the meeting of the American Educational Research Association, New Orleans, 1973.

Cohen, J. Statistical power analysis for behavioral sciences. New York: Academic Press, 1969.

Commission on Instructional Technology. Report to the President and Congress: To improve learning. Washington, D. C.: United States Government Printing Office, 1969. (ERIC ED 043 240)

Cook, M. J. An experimental research design for comparing the retention effects of different instructional treatments. Baltimore: University of Maryland, 1970. (ERIC ED 040 578)

Corrozzi, J. F. The effects of reading time, type of question, and instructional format on short and long-term retention of relevant and incidental prose material. (Doctoral dissertation, University of Delaware) Ann Arbor: University Microfilms, 1970, No. 71-6465.

Cunningham, D. J. The retention of connected discourse: A review. Review of Educational Research, 1972, 42, (1), 47-71.

Davis, R. A. Learning in the schools. Belmont, Calif.: Wadsworth, 1966.

Deady, G. M. The effects of an increased time allotment on student attitudes achievement in science. Chico State College, Calif., 1970. (ERIC 039 136)

Denny, R. R. A two-year study of the effects of an increased time allotment upon achievement in arithmetic in intermediate grades: Field study #1. Greely, Colo.: Colorado State College, 1955.

Deterline, W. A. Educational systems. In Technology and innovation in education. New York: Praeger, 1968.

Dick, W., Merrill, P. & O'Neal, H. Systems approach model for the development of instructional materials. Tallahassee: Computer Assisted Instruction Center, Florida State University, 1970.

Dixon, W. J. (Ed.) BMD: Biomedical computer programs. Berkeley: University of California Press, 1970.

Draper, N. R. & Smith, H. Applied regression analysis. New York: John Wiley, 1966.

Dyer, H. S. An educational researcher's view of systems analysis. Educational Technology, September, 1969, 83-85.

Fenton, E. The new social studies. New York: Holt, Rinehart & Winston, 1967.

Frase, L. T. Cybernetic control of memory while reading connected discourse. Journal of Educational Psychology, 1969, 60, 49-55.

Freud, S. The psychopathology of everyday life. New York: W. W. Norton, 1965.

Gagne, R. M. & Bassler, O. C. Study of retention on topics elementary non-metric geometry. Journal of Educational Psychology, 1963, 54, (3), 123-131.

Gagne, R. M. & Rohwer, W. D. Instructional psychology. Annual Review of Psychology, 20, 1969, 381-418.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart & Winston, 1970.

Gagne, R. M. The domains of learning. Presidential address, American Educational Research Association, New York, 1971.

Gay, L. R. Temporal position of reviews and its effect on retention of mathematical rules. (Doctoral dissertation, Florida State University) Ann Arbor: University Microfilms, 1971, 71-239.

Gibson, J. R. Transfer effects of practice variety in principle learning. Unpublished doctoral dissertation, University of California, Berkeley, 1969.

Glaser, R. The design of instruction. In M. D. Merrill (Ed.) Instructional design: Readings. Englewood Cliffs, N. J.: Prentice Hall, 1971.

Glaser, R. Teaching machines and programmed learning II. Washington, D. C.: National Education Association, 1965.

Governor's Committee on Public Education. The report of the Governor's committee on public education in Texas. Austin: Texas State Department of Education, 1970.

Hall, J. F. Verbal learning and retention. New York: J. B. Lippincott, 1971.

Hebb, D. O. & Foord, E. N. Errors of visual recognition and the nature of trace. Journal of Experimental Psychology. 1945, 35, 235-248.

Hull, C. L. A behavior system. New Haven: Yale University Press, 1952.

Jarvis, O. T. Time allotments and pupil achievement in the intermediate grades. Houston; Texas Bureau of Educational Research and Service, 1962. (ERIC ED 035 063)

Kimble, D. P. (Ed.) The organization of recall. New York: New York Academy of Sciences and Interdisciplinary Communications Program, 1964.

Kimble, G. A. Conditions and learning. New York: Appleton Century-Crofts, 1961.

Koffka, D. Principles of Gestalt psychology. New York: Harcourt Brace, 1935.

Kohler, W. Gestalt psychology. New York: Liveright, 1947.

Krumboltz, J. Factors affecting the design of effective teaching machine programs. 1964. (ERIC ED 026 853)

Lumsdaine, A. A. Instructional research: Some aspects of its status. Journal of Experimental Education, Fall, 1968, 95-101.

Lysaught, J. P. & Williams, C. M. A guide to programmed instruction. New York: Wiley, 1963.

Massialas, B. G. & Smith, F. R. (Eds.) New challenges in social studies. Belmont, Calif.: Wadsworth, 1965.

Mowrer, G.E. A study of the effect of length of time of high school english periods in achievement in English. Unpublished doctoral dissertation, University of Missouri, 1956.

McGeogh, J. A. Forgetting and the law of disuse. Psychological Review, 1932, 39, 352-370.

Osgood, C. E. Method and memory in experimental psychology. New York: Oxford Press, 1953.

Penfield, W. Neurophysical basis of higher functions of the nervous systems. In J. Field (Ed.), Handbook of physiology. Washington, D. C.: American Physiological Society, 1960.

Postman, L. Retention as a function of overlearning. Science, 1962, 135, 666-667.

Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.

Reynolds, J. H. & Glaser, R. Effects of repetition and spaced review upon retention of a complex learning task. Journal of Educational Psychology, 1964, 55, 297-308.

Schure, A. Science education and instructional systems. In R. A. Weisgerber (Ed.), Instructional process and media innovation. Chicago: Rand McNally, 1968.

Seagoe, M. V. The learning process and school practice. Scranton, Pa.: Chandler Publishing, 1970.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Spence, K. W. Mathematical theories of learning. Journal of General Psychology, 1952, 49, 283-391.

State Committee on Goals for Education. Goals for education in Georgia. Atlanta: Georgia State Department of Education, 1970.

State Department of Education. Goals for education in Florida. Tallahassee: Florida State Department of Education, 1964.

Sulzen, R. H. The effects of empirical program revision and presentation of objectives on student performances. Denver: Air Force Human Resources Laboratory, 1972.

Thorndike, E. L. Educational psychology. New York: Teachers College Press, 1913.

Toulmin, S. E. The uses of argument. Cambridge: Cambridge University Press, 1958.

Wang, M. & Lindvall, C. M. An exploratory investigation of the Carroll learning model and the Bloom strategy for mastery learning. Pittsburgh: University of Pittsburgh Learning, Research, and Development Center, 1969. (ERIC ED 028 841)

Welch, W. W. & Bridgham, R. G. Physics achievement gain as a function of teaching duration. School Science and Math, May, 1968, 449-454.

Wickelgren, W. A. Acoustic similarity and retroactive interference in short term memory. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 53-61.

Wilson, K. V. A distribution-free test of analysis of variance hypotheses. Psychological Bulletin, 53, 1, 1956, 96-101.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Winne, P. H., Houck, W. E., & Moore, J. W. Acquisition and retention as a function of three types of repetition. Paper presented at the meeting of the American Educational Research Association, Chicago, 1972.

Wodtke, K. H. On the assessment of retention effects in educational experiments. Journal of Experimental Education, Summer, 1967, 28-36.