R E P O R T R E S U M E S

ED 013 256 TE 000 017

THE EFFECTIVENESS OF PROGRAMED MATERIALS IN ENGLISH SYNTAX AND THE RELATIONSHIP OF SELECTED VARIABLES TO THE LEARNING CONCEPTS.

BY- BLOUNT, NATHAN S. AND OTHERS
WISCONSIN UNIV., MADISON

REPORT NUMBER WU-TR-17 PUB DATE JAN 67
REPORT NUMBER BR-5-0216-TR-17
CONTRACT OEC-5-10-154
EDRS PRICE MF-$0.50 HC-$2.52 63P.

DESCRIPTORS- *ENGLISH INSTRUCTION, *PROGRAM EVALUATION, *PROGRAMED MATERIALS, *RESEARCH AND DEVELOPMENT CENTERS, *SYNTAX, FEEDBACK, GRAMMAR, INSTRUCTIONAL MATERIALS, INTELLIGENCE QUOTIENT, JUNIOR HIGH SCHOOL STUDENTS, PROGRAMED INSTRUCTION, REVIEW (REEXAMINATION), TEACHING TECHNIQUES, TRANSFORMATION THEORY (LANGUAGE), WRITING EXERCISES, UNIVERSITY OF WISCONSIN, MADISON, R D CTR. FOR LEARNING AND RE-EDUC.,

THE RESEARCH AND DEVELOPMENT CENTER FOR LEARNING AND RE-EDUCATION AT THE UNIVERSITY OF WISCONSIN CONDUCTED AN EXPERIMENT IN THE SUMMER OF 1966 TO TEST AND IMPROVE ITS PROGRAMED INSTRUCTIONAL MATERIALS IN STRUCTURAL AND TRANSFORMATIONAL GRAMMAR AND TO EXTEND KNOWLEDGE ABOUT CERTAIN VARIABLES RELATED TO EFFICIENCY OF LEARNING CONCEPTS--THE IQ LEVELS AND SEX OF STUDENTS, AND THE EFFECTIVENESS OF ADVANCED ORGANIZERS, REVIEW OF MATERIAL, NEGATIVE INSTANCES OF CONCEPTS, INFORMATIVE FEEDBACK, AND WRITING EXERCISES. FOR EACH PROGRAMED LESSON COMPLETED BY THE 48 PRE-EIGHTH GRADERS, TIME FOR COMPLETION AND PERCENTAGE OF ERRORS WERE RECORDED. RESULTS SHOWED THAT THE AVERAGE COMPLETION TIME OF A LESSON WAS WELL WITHIN THE LIMITS OF A CLASS PERIOD AND THAT, WITH AN ERROR RATE OF 7.3 PERCENT PER LESSON, THE MATERIAL WAS EFFECTIVELY LEARNED. MULTIPLE CHOICE AND COMPLETION TESTS MEASURED LEARNING PROGRESS. SEX AND IQ SHOWED HIGH AND CONSISTENT CORRELATIONS--FEMALES TENDED TO SCORE HIGHER ON TESTS THAN MALES, AND TEST RESULTS CLOSELY PARALLELED IQ LEVELS. THOSE STUDENTS WHO RECEIVED INFORMATIVE FEEDBACK AND WRITTEN EXERCISES SCORED CONSISTENTLY HIGHER ON TESTS THAN THOSE WHO DID NOT RECEIVE THESE AIDS. THEREFORE, INCORPORATION OF THESE TECHNIQUES IN THE USE OF THE PROGRAMED MATERIAL IN THE JUNIOR HIGH SCHOOL SEEMS ADVISABLE. ON THE OTHER HAND, ADVANCED ORGANIZERS, REVIEWS, AND NEGATIVE INSTANCES OF CONCEPTS DID NOT CORRELATE SIGNIFICANTLY WITH TEST SCORES. (DL)

THE EFFECTIVENESS OF PROGRAMED MATERIALS IN ENGLISH SYNTAX AND THE RELATIONSHIP OF SELECTED VARIABLES TO THE LEARNING OF CONCEPTS

RESEARCH AND DEVELOPMENT CENTER FOR LEARNING AND RE-EDUCATION

Technical Report No. 17

THE EFFECTIVENESS OF PROGRAMED MATERIALS IN ENGLISH SYNTAX AND THE RELATIONSHIP OF SELECTED VARIABLES TO THE LEARNING OF CONCEPTS

Nathan S. Blount, Herbert J. Klausmeier, Shelby L. Johnson,
Wayne C. Fredrick, and James G. Ramsay

Research and Development Center
for Learning and Re-education
The University of Wisconsin
Madison, Wisconsin

January 1967

The research and development reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program.

Center No. C-03 / Contract OE 5-10-154

PREFACE

The R & D Center for Learning and Re-education has as its primary goal the improvement of cognitive learning in children and youth, commensurate with good personality development. Through synthesizing present knowledge and conducting research to generate new knowledge, we are extending the understanding of cognitive learning and the conditions associated with efficiency of school learning. Knowledge is being focused upon the three main problem areas of the Center: developing exemplary instructional systems, refining the science of human behavior and learning as well as the technology of instruction, and inventing new models for school experimentation, development activities, and so on.

The design and conduct of the study described in this Technical Report was a cooperative effort of groups representing two of the Center's problem areas. Primary responsibility for the development and field testing of the instructional materials and for related experiments was assumed by Professor Blount and Mrs. Johnson, whose activities in the Center are directed toward the development of an exemplary instructional system in English syntax. Two of my research assistants in concept learning, Mr. Fredrick and Mr. Ramsay, and I assumed primary responsibility for designing the experiments dealing with variables relevant to concept learning in the school. The joint activity of researchers in the areas of learning theory and curriculum has been planned cooperatively from the outset. Following the experimentation during the summer of 1966, a third version of the programed materials was prepared by Mr. Blount and Mrs. Johnson for further field testing and an evaluation of the effects of the programed instruction on writing abilities of eighth graders.

Herbert J. Klausmeier
Co-Director for Research

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In this technical report are given the results of: (1) field testing and possibly improving instructional materials in structural and transformational grammar through analysis of (a) the students' comments on individual frames of the material, (b) error rates, (c) the varying amounts of time required by students to complete the lessons, and (d) students' absolute gains from pretest to posttest; and (2) experiments to extend knowledge about variables related to efficiency of learning concepts, the more specific variables being sex of the students, IQ level of the students, advance organizer, review of material, negative instances of concepts, informative feedback, and writing exercises.

Forty-eight Ss, pre-eighth graders in Summer 1966, were used for purposes of statistical analysis of the experiments. In the experiments, factorial designs were used; sex of the students and level of IQ were stratifying variables while the other variables were manipulated. The dependent variables used to measure effectiveness of the 21 lessons of programed material were (1) time per lesson and (2) errors per lesson. The dependent variables in the experiments were scores from a total of 441 multiple-choice test items and 338 completion items presented to the Ss over the 24 day treatment.

Analysis of Ss' performance on the programed material showed an average time of 21 minutes required to complete each lesson and an overall error rate of 7.3 per cent per lesson. Sex and IQ showed highly consistent effects: on every test, the High IQ group scored highest, the Low group scored lowest, and the Medium group scored between the High and the Low groups. The results from the experiments also indicated that females tended to score higher on the tests than males. Feedback was a significant factor, the feedback group scoring consistently higher than the no-feedback group. Ss who received writing exercises scored significantly higher ($p < .01$) than Ss who did not receive writing exercises. The variables advance organizer, review, and negative instances did not reach statistical significance.

I

INTRODUCTION

The studies which are presented in this paper report experimentation conducted in the summer of 1966 (1) to assess the effectiveness of instructional materials presenting selected concepts from English syntax and (2) to investigate the variables sex, IQ level, advance organizer, review, negative instances of concepts, informative feedback, and writing exercises as these variables affect the learning of the instructional materials.

BACKGROUND STUDIES AND RATIONALE FOR DEVELOPING THE INSTRUCTIONAL MATERIALS IN ENGLISH SYNTAX

Various correlational studies of the relationship of grammar and writing ability conducted since the turn of the century by such researchers as Asker (1923), Boraas (1917), Bradford (1941), Catherwood (1932), Hoyt (1906), Rapeer (1913), Robinson (1960), and Segal and Barr (1926) have suggested that traditional grammar as it has been taught has had little effect on the improvement of writing skills and that traditional grammar has had little if any transfer value in developing writing ability. Boraas (1917), for example, found a lower correlation between knowledge of traditional grammar and ability in composition than between knowledge of traditional grammar and knowledge of arithmetic and history.

However, the studies cited above examined knowledge of traditional grammar in relationship to ability in written composition. And the last several decades have seen the emergence of several new grammars of English: structural grammar, arising largely from the work of C. C. Fries at the University of Michigan; and transformational grammar, arising from the work of Noam Chomsky at MIT. Experimental evidence is needed to show the contributions of the new grammars to improving writing.

The traditional school grammar is often prescriptive. It emphasizes somewhat arbitrary rules rather than providing descriptions of the language. It pays little attention to word order or to the patterns of English sentences. Its definitions of parts of speech sometimes rely more on meaning than on form. It has been criticized for relying on the apparatus of Latin grammar and on models of eighteenth century grammars of English.

Structural grammar, conversely, is largely the product of studying human linguistic behavior scientifically. Observing language in action inductively, the structural linguist has analyzed the spoken language before examining written language. In classifying words as a member of a form class, he has looked at such devices of form as word order, function-words, inflections, and formal contrasts as signals of structural meaning.

Transformational, or generative, grammar employs a deductive process by which it tests or verifies various generalizations or hypotheses about language. It offers a theory of kernel sentences and a theory of intuition by which native speakers of English gain an understanding of the grammar of their language. It seeks to set forth a list of rules by which English words can be combined to produce every English sentence. Transformational grammar presents rules which enable a native speaker to produce from a kernel sentence such as "Bob broke the window" the transformations, "Did Bob break the window," "Bob didn't break the window," "The window was broken by Bob," "which Bob broke," and "who broke the window."*

Scholars formulating systematic analyses of language have often been unwilling to speculate

---

*The details of these grammars are, of course, beyond the scope of this work. For a selective bibliography which lists simply written, as well as highly technical, descriptions of recent work in linguistics see O. Thomas, Transformational Grammar and the Teacher of English (New York: Holt, Rinehart and Winston, Inc., 1965), pp. 226-230.

on the uses of the new grammars in the public schools. However, a few studies have been conducted which examine the values of structural and of transformational grammars for the teaching of writing.

Three investigators—Blake (1964), Johnson (1960), and O'Donnell (1963)—found no evidence of superiority of a structural approach to a traditional grammar approach in teaching writing. O'Donnell (1963), using correlation analysis, found no basis for assuming that either knowledge of traditional grammar or knowledge of structural grammar would be regularly accompanied by excellence in written composition. A study by Lin (1965) failed to reveal any statistically significant differences between an experimental group receiving pattern practice (relying on structural grammar) and a control group.

To date, only one researcher has reported experimentation with teaching transformational grammar in the secondary schools. A two-year experiment in transformational grammar supported by the U. S. Office of Education (Project Number 1746), reported by Zidonis (1965), one of the major investigators, sought to answer questions as to (1) whether ninth- and tenth-grade students could apply the transformational rules of generative grammar in their own writing, (2) whether the proportion of well-formed sentences in student writing might increase over a two-year period, (3) whether knowledge of transformational grammar might increase the students' repertoire of grammatical structures, and (4) what transformational and co-occurrence errors might occur in the compositions, and to what extent such errors might increase or decrease during a two-year period.

At the beginning of the study, 50 ninth-grade students were assigned at random to two classes taught by two teachers also assigned at random. For two years, both classes studied the same content except that one was taught concepts and principles of generative grammar. The experimental class (generative grammar) studied phrase structure rules the first year and transformational rules the second year. Writing collected from both experimental and control groups early in the first year and late in the second year exceeded 70,000 words. The instrument used to analyze the sentences in the sample contained three component measures: (1) proportion of well-formed sentences, (2) structural complexity score, and (3) error change score. Comparisons between experimental and control classes were made by analysis of variance.

Data on proportion of well-formed sentences showed a difference ($p < .01$) between the experimental and the control classes and a difference ($p < .01$) between pretest and posttreatment scores. The investigators attributed the gains in score by the experimental class to the study of transformational grammar.

Data on structural complexity scores showed no significant difference between the experimental and the control groups. Both classes increased their structural complexity scores ($p < .01$).

Data on error change scores indicated two categories of error, misapplication of a transformational operation and co-occurrence error, to be of interest.

The research supported the assumptions that ninth- and tenth-grade students can master the concepts of transformational grammar fairly easily, that knowledge of generative grammar increases significantly the proportion of well-formed sentences, and that knowledge of generative grammar reduces the occurrence of certain errors in composition.

Further information for the researcher interested in the relationship of a knowledge of grammar to writing ability is to be found in a series of studies by Hunt (1964, 1965, 1966).

One of Hunt's recent reports (1965) extends his earlier work on objective indexes of maturity in writing. Analyzing 1,000-word samples of writing by 18 superior adults (publishing expository articles in Harper's and Atlantic), and comparing these analyses with previous analyses of writing by 18 fourth-grade, 18 eighth-grade, and 18 twelfth-grade students, Hunt reported T-unit length, length of minimal terminable unit (one main clause plus any clauses attached to or embedded in it), to be a more useful tool in providing an index of level of maturity in writing than sentence length, subordination ratio, or mean clause length. One of the findings of the analysis of prose written by superior adults was that an increase in T-unit length produces an average sentence length 47 per cent above that for twelfth-grade students. Another finding was that the average length of clauses in adult prose is 36 per cent greater than that for twelfth-grade students. A third finding was that the average length of T-units is 40 per cent greater for the adults than for twelfth graders. More than any other single factor, clause length distinguishes between superior adults and twelfth-grade students of average mental ability.

In the same publication, Hunt (1965) discusses the feasibility of producing a sentence-building program for students in the early years of the secondary school. Such a program would

reduce independent clauses to subordinate clauses and to nonclauses; it would help students consolidate subordinate clauses and nonclauses with adjoining clauses and with T-units. Hunt suggests that such a program might help the student "work up to structures of considerable depth and complexity comparable to those exhibited by twelfth graders and superior adults [p. 157]." He further suggests that such a sentence-building program might be incorporated into a course in transformational grammar.

Thus, data supplied by Hunt suggest that mean length of T-units, mean clause length, ratio of subordinate to all clauses, ratio of clauses per T-unit, and mean sentence length are statistically significant indexes of maturity in writing, discriminating between the writing of eighth graders, twelfth graders, and superior adults. Data supplied by Zidonis suggest that students in the junior high school can master the concepts of transformational grammar and that such concepts transfer to writing.

OTHER OBJECTIVES FOR THE PROJECT IN ENGLISH SYNTAX

Long-range objectives include identifying concepts from recent scholarship in English grammar which might help young people gain insight into English syntax as it appears in certain declarative sentences called basic sentence patterns and in several transformations of these basic sentences. Selected concepts have been incorporated into instructional materials which are programed in linear style. The effectiveness of these materials is to be assessed in terms of error rate and time involved. In addition to attempting to make the customary assessment of the effectiveness of the programed materials in terms of error rate, etc., the investigators have as a main objective the assessment of the effect of the study of concepts from structural and transformational grammars on student writing. To study the effect of the instructional materials on writing, the investigators are to analyze 1,000-word themes written by eighth-grade students pre- and posttreatment, both for students in control and treatment groups. The analyses of writing will examine five factors—mean length of T-unit, mean clause length, ratio of subordinate to all clauses, ratio of clauses per T-unit, mean sentence length—factors which Hunt's research has shown statistically significant indexes of maturity of writing. Other objectives include field-testing the programed materials large-scale and using the materials as an experimental instrument by which variables of interest in concept learning can be examined in a school setting.

CONTENT OF THE MATERIALS IN ENGLISH SYNTAX

Prompted by ideas suggested by the work of Hunt and of Zidonis, work on the instructional materials was begun in September 1965, when Blount and Johnson began to build a program designed to teach structural and transformational grammars. The main concepts which were chosen and included in the program are listed in the following outline.

General Outline of Concepts in Program

BASIC SENTENCE PATTERNS

Lesson 1: A Look at Basic Sentences
- --basic sentence
- --subject and predicate groups
- --NP
- --noun test-sentence

Lesson 2: Basic Sentence Patterns 1 and 2
- --completer
- --adj test-sentence

Lesson 3: Basic Sentence Pattern 3 and Review
- --adv test-sentence
- --prep phrase

Lesson 4: Basic Sentence Pattern 4
- --VP
- --verb (identify by position)
- --optional adverb

Lesson 5: More About Nouns
- --noun marker
- --inflectional endings

Lesson 6: Recognizing Verbs
- --inflectional endings
- --verb marker
- --past$_1$ and past$_2$ forms

Lesson 7: Recognizing Adverbs
- --inflectional endings
- --affix
- --markers

Lesson 8: Recognizing adjectives with Patterns 5 and 6
- --Pattern 5: Vs
- --Pattern 6: Vb
- --adj: marker
  inflection
  affix

Lesson 9: Basic Sentence Patterns 7 and 8
- --two-word verbs
- --pronouns as objects

Lesson 10: Basic Sentence Pattern 9 and General Review

TRANSFORM PROCESSES

Lesson 11: T-rel (transform process for constructing related clauses)
--related clause
--relating pronoun
--transform process
--insert and base

Lesson 12: T-rel Again
--relating adverb

Lesson 13: T-rel after T-pos (transform process for possessive construction)

Lesson 14: T-pass (transform process for passive construction)

Lesson 15: T-ph (transform process for constructing participial and prep phrases)
-- -ing phrase
--$past_2$ phrase
--prep phrase

Lesson 16: T-BN (transform process for placing single-word modifiers before the noun)

Lesson 17: T-NP and double nouns (transform process for constructing appositives)
--double nouns (noun adjunct)

Lesson 18: T-NC (transform process for constructing noun clauses)
--noun clause

Lesson 19: T-V_{to} (transform process for constructing infinitives as nominals)
--nominal
--V_{to}
--V_{to} phrase

Lesson 20: T-V_{ing} (transform process for constructing -ing nominals)
--V_{ing}
--V_{ing} phrase

Lesson 21: Adj & Adv Nominals and Review of Nominals

Thus, Lessons 1-10 contain the concepts of the four form classes noun, adjective, verb, and adverb together with discussion of the several attributes, several signals of syntactic structure, of each form word such as word order, inflections, and function words. The presentation of the form classes follows the work of such structural grammarians as Francis (1958), Fries (1940), Roberts (1956), and Whitehall (1951). Lessons 1-10 also present basic sentence patterns out of which longer or more complex sentences are formed through rearrangement, addition, deletion, or combinations of elements of basic sentences. The basic sentence patterns discussed in the instructional materials are:

$$\left.\begin{array}{l}
\text{Pattern 1: NP + be + NP} \\
\text{Pattern 2: NP + be + adj} \\
\text{Pattern 3: NP + be + adv} \\
\text{Pattern 4: NP + V} \\
\text{Pattern 5: NP + } V_s \text{ + adj} \\
\text{Pattern 6: NP + } V_b \text{ + } \begin{Bmatrix} \text{adj} \\ \text{NP} \end{Bmatrix} \\
\text{Pattern 7: NP + V + NP} \\
\text{Pattern 8: NP + V + NP + NP} \\
\text{Pattern 9: NP + V + NP + } \begin{Bmatrix} \text{adj} \\ \text{NP} \end{Bmatrix}
\end{array}\right] \rightarrow \text{+ (adv)}$$

Lessons 11-21 treat transformations of basic sentences. Included are the transform process for constructing related clauses, the transform process for constructing noun clauses, the transform process for constructing -ing nominals, and so on. The concepts presented in Lessons 11-21 rely heavily on the work of such scholars as Bach (1964), Chomsky (1964), Lees (1966), Roberts (1964), Rogovin (1964), and Thomas (1965).

THE EXPERIMENTAL VARIABLES

The experimental variables which were incorporated into the instructional materials were: advance organizer, review of material, negative instances of concepts, informative feedback, and writing exercises.

Advance Organizers

Frequently, tasks to be learned by subjects in the laboratory are of limited duration and the materials are of limited meaningfulness, while in the school setting, several weeks or months may be needed for mastery of a particular skill and, in the latter case, the learner may be exposed to much greater amounts of meaningful material than would be presented in an hour's session in the laboratory. Ausubel (1963) has suggested that the learning of meaningful material may be facilitated through the use of advance organizers.

Ausubel characterizes the learner as entering the learning situation with an existing cognitive structure. If the cognitive structure is adequate with respect to new material to be learned, initial learning of the new material is facilitated in that the new material may be subsumed into the structure with relative ease and efficiency. If the structure is inadequate, learning of the new material will be more difficult.

One way to help insure an adequate cognitive structure is, according to Ausubel, through the use of advance organizers. Advance organizers introduce the material to be learned, but at a higher level of abstractness and generality than the material itself. They provide the higher-order concepts, the conceptual framework into which new material can be subsumed.

Because of the nature and amount of material to be covered by the Ss in the English Project, advance organizers seemed to be an appropriate way to help Ss acquire and retain the information presented to them. Because of the programed format of the English materials, advance organizers were incorporated into an experimental design in order to test their effectiveness on material presented frame by frame in a linear sequence.

Review

Just what amount of repetition and review a certain sequence of programed instruction requires is difficult to know. Glaser (1965) mentions a study in which one-fifth of the program frames were removed, with the result being no appreciable loss of learning. Greeno (1964) found that the many repetition frames of early programs had little value in increasing retention. Reynolds and Glaser (1964) also found little effect due to amount of repetition but did find that a spaced review sequence facilitated retention of the reviewed material. In the present program, an attempt was made to determine whether additional spaced review was needed over and above the amount already built into the sequence of frames.

Negative Instances

For Lessons 11-15, negative instances of concepts replaced one-fourth of the frames, which were normally positive instances. The use of the negative instances was an attempt to apply and test conclusions from laboratory studies. These studies, involving concept identification tasks which are rather unique as compared to school learning, have shown, in general, that initial learning is more efficient when positive instances of the concept rather than negative or a mixture of positive and negative instances are presented. Hovland and Weiss (1953) have shown that performance is best with positive instances only and poorest with negative only, while the mixture of the two is intermediate. Smoke (1932) determined that negative instances were "not necessarily aids to learning," but this conclusion was revised since it appears that negative instances discourage hasty judgments (Smoke, 1933). Adding negative instances caused no improvement and seemed to slow down concept attainment in a study by Mayzner (1962), but adding positive instances was beneficial. Other experimenters have found only slight differences in concept tasks between positive instances conditions and mixed instances (Olson, 1963; Smoke, 1932, 1933). The problem seems to be that subjects overlook or ignore information from negative instances (Braley, 1963; Wason, 1959). Another factor suggested is that reasoning from negative instances may be intrinsically more difficult than reasoning from positive instances (Huttenlocker, 1962; Wason, 1959). A third point is that negative instances usually contain less information than positive (Hovland & Weiss, 1953).

Negative instances do, however, play a role in concept formation, as Smoke has suggested above. Buss (1950) showed that negative instances restrict the range of values to which the concept or class response will be made. Negative instances may also act as a "contrast class" by which dimensions and values of the concept are defined (Wallach, 1958).

It seemed reasonable to ask whether the addition of negative instances to a program of grammar would have any noteworthy effects.

Informative Feedback

In a survey of the literature on programed learning for the school years 1961-1962 and 1962-1963, Hanson and Komoski (1965) indicated that there was a decreasing trend for teachers simply to serve as proctors when using programed material. Two valuable functions of the teacher employing programed material are (1) providing feedback as to how the student is performing and (2) providing positive reinforcement. A treatment which included direct teacher involvement was thus incorporated in the design of one of the experiments in the present study.

Writing Exercises

Research conducted in the last several decades has suggested that practice in writing does not improve ability in writing. However, composition has been evaluated considering a variety of factors such as spelling, punctuation, style, and so on. Few studies have examined the impact of recent grammars of the structure of the single sentence on syntactic maturity, nor have refined techniques of analy-

sis of syntactic maturity been available prior to the studies directed by Kellogg Hunt. As an exploratory study, it was thought suitable (1) to consider the effect of a study of the concepts from recent grammars on sentence-combining skills; (2) to make judgments of such an effect by analyzing briefly the use of certain syntactic structures within the sentence, disregarding punctuation, spelling, paragraph development and so on; and (3) to attempt to develop skills in sentence-combining transforms through structured writing exercises.

Grammar textbooks have traditionally included various kinds of exercises mainly on an intuitive basis since no research has supported the idea that isolated practice in a particular skill will improve writing ability. Feeling that some guidance might be needed in transferring the concepts presented in the program on the processes involved in understanding and manipulating sentence structures, the investigators constructed six writing exercises to follow Lessons 16-21. Each exercise consisted of several sections, some of which were highly structured, others of which were structured only slightly to allow Ss some degree of freedom in constructing their responses. In order to place equal emphasis on the different processes involved in writing and understanding more mature sentences, one sentence of each exercise was intended to give practice in at least one of the following skills: (1) creating original structures of the kind called for in the writing exercise instructions and placing the structures within the framework of a sentence; (2) tracing the structures in a transform sentence back to the basic sentences from which they derive; and (3) building one longer transform sentence from a series of three to eight basic sentences through the processes of deleting and embedding.

OBJECTIVES OF THE SUMMER PROGRAM, 1966

Before any studies could be made of the impact of the instructional materials on student writing in terms of analyses of substantial amounts of writing, it was necessary to field test the materials on a small scale. Such field testing to provide data for the revision of the program was one objective of the summer program. Because the material appeared to have qualities which would make it a useful experimental instrument in its present form, study of several learning variables applicable to a school setting were incorporated into the summer project.

More specifically, the objectives of the summer program included:

1. Field testing and possibly improving the instructional material through analysis of (a) the students' comments on individual frames of the material, (b) error rates, (c) the varying amounts of time required by students to complete the lessons, and (d) students' absolute gains from pretest to posttest.

2. Conducting experiments to extend knowledge about variables that might be related to efficiency of learning concepts, the more specific variables being sex, IQ level, advance organizer, review of material, negative instances of concepts, informative feedback, and writing exercises.

Inasmuch as information concerning the last five variables in (2) above can be incorporated into the programed materials, the experiments dealing with these variables might also contribute to revision or improving the programed material.

II

METHOD

SUBJECTS

Fifty-nine pre-eighth graders participated in the Summer English Program. Data from 48 of these students were used for statistical analysis. Sex and IQ were used as stratifying variables in the experimental designs. The 48 Ss included 24 boys and 24 girls. Eight boys and eight girls were in a low IQ range relatively (90-107); eight boys and eight girls in a middle range (113-120); and eight boys and eight girls in a high range (126-140).

The students who participated in the program were selected from two schools in Madison, Wisconsin: Van Hise Junior High School and Cherokee Heights School. Class lists from the two schools were obtained of those students who had been in seventh grade in the 1965-66 school year. This population consisted of 504 students. Composite IQ scores from the L and NL parts of the California Test of Mental Maturity (CTMM) were obtained for each student in the population. The test had been administered when the students were in fifth grade. In relation to national norms, the distribution of scores for this population was extremely high, the modal score being 126. From the population, students in three ranges of scores were identified: those students who scored between 90 and 107; those who scored between 113 and 120; and those who scored between 125 and 140. These ranges were labeled "low," "medium," and "high," respectively. The standard error of measurement on the CTMM at the fifth-grade level is five, so the adjacent ranges were separated by one standard error of measurement.

Sampling began with the students in the low range. Letters inviting the student to participate in the program were sent to the parents of a random sample of 50 of the 76 students in the low range. The letter included a postcard to be returned which would indicate whether or not the student was interested in participation. These letters, and those sent to the parents of students in the middle and high ranges, informed the parents that students participating in the program would be paid one dollar an hour for their participation. Twenty-four of the students volunteered from the low group, and ten boys and ten girls were selected randomly for participation.

Because of the high number of students in the low range who volunteered to participate, it was decided to send letters and postcards to a smaller proportion of the students in the medium and high ranges. A random sample of 32 of the 119 students in the medium and high range was contacted. Twenty students volunteered, and nine boys and nine girls were randomly selected. Of the 148 students in the high range, 32 were contacted; 21 volunteered, and nine boys and nine girls were randomly selected. Because of the possibility that certain experimental treatments might lead to drop-outs from the program, three additional students were contacted and participated in the experiment. The final distribution of subjects is presented in Table 1.

Table 1

Distribution of Ss

	IQ Range		
	Low (90-107)	Medium (113-120)	High (126-140)
Males	10	9	9
Females	10	10	11

Since there were no drop-outs, eight Ss from each cell were randomly selected, and the data from these Ss were used for statistical analysis of the experiments.

SETTING, MATERIALS, AND PROCEDURES

The Summer English Project took place in the chorus room of Van Hise Junior High School, Madison, Wisconsin. The room had a capacity of 70 and was equipped with desks.

The program consisted of 21 lessons, with each lesson averaging about 65 frames. The range in number of frames was from 41 on the shortest lesson to 85 on the longest. One lesson was given to the Ss each day. Subjects were prepared to participate for two hours each morning (8:30-10:30), Monday through Friday, from June 20 to July 22. An experimenter and a supervisor administered the program, tests, and experimental treatments.

The instructional materials, in 21 lessons, were programed in linear style. The program was constructed with the expectation of a low error rate and required an overt, written response to each frame either by constructing a response or by selecting a response from alternatives. The attempt was made to make each student's response as nearly without error as possible as the result, in part, of various cueing techniques such as boldface type. The attempt was also made to write the programed instructional materials in such a way that the student need receive little or no additional assistance. As well as illustrating negative instances, Appendix B of this report, Lesson 15 of the instructional program, presents frames from the programed materials as an example of the style of programing.

Ss responded to question frames located on the left column of the lesson page by writing their answer in ink on a lesson answer sheet. They then moved a cardboard shield down the right column of the lesson page until they uncovered the answer to the frame to which they had just responded. If their answer was correct, they went on to the next question frame; if not, they crossed out their wrong answer and wrote in the correct one above it. Space was also provided after each answer on the answer sheet for Ss to write comments about the frame.

During a typical daily session, Ss began work on that day's lesson at 8:35; the last S had usually finished by 9:20; Ss who finished early read library books or magazines until all Ss had completed the lesson. Each S was thus allowed to work at his own rate, while the constant rate of one lesson per day was still maintained. When the program answer sheets had been collected, the Ss were given a five-minute break. The test over the day's lesson was then administered. The last S had usually finished by 9:55, and the Ss were usually dismissed around 10:00.

Experiment I

Experiment 1 was conducted during the first 10 lessons of the program and was designed to test the effect of advance organizers. One advance organizer was presented before each of the first 10 lessons to half of the Ss. The other half of the Ss read control material of equivalent length. The advance organizer and control material for Lesson 1 have been included in Appendix A of this report.

Experiment 2

Experiment 2 took place during the tenth lesson and was designed to test the effect of review as opposed to no review on the first 10 lessons. Lesson 10 introduced the final basic sentence pattern and was then devoted to review frames covering all of the basic sentence patterns and the important terminology which the Ss had encountered during the first 10 lessons. The review material was in a programed format as were the other lessons.

Experiment 3

Experiment 3 was administered during Lessons 11-15 and tested (1) the effects of negative instances of concepts in the program and (2) the effects of feedback of daily test results.

Negative instances of concepts Half of the Ss in the feedback group and in the no feedback group received negative instances as a treatment variable. The negative instances were examples of what the concepts in the lessons were not. As an example, two negative frames from Lesson 15 were:

3. < My mother gave that man a book. >
   Is that man a book a related clause?
   Ans. No
24. T-ph does not produce past___ phrases.
   Ans. 1

For each of Lessons 1 through 15, one-fourth of the frames were deleted and a negative frame was put in its place. In this manner, both treatments were kept the same in length and format with the only difference being that 14 to 20 of the usual frames were replaced with the negative examples and statements. Appendix B contains Lesson 15 as a sample of a lesson containing negative instances.

Effects of feedback During these lessons, half of the Ss met in an adjacent classroom where the previous day's tests were distributed. On these tests, wrong responses were marked, correct responses to the items missed were written in, and positive comments commending the Ss' performances on the test were written

on the last page of all the tests. Before this group of Ss began on the daily lesson, the experimenter went through the corrected test with the Ss, concentrating on those items which had been missed most frequently. Any questions which Ss raised were answered during this session. This combination of feedback of test results, positive comments, and discussion was felt to be analogous to a classroom situation in which the teacher provides feedback and positive reinforcement.

Experiment 4

Experiment 4 took place during Lessons 16-21 and tested the effect of Ss doing writing exercises relevant to that day's lesson before taking the daily test. Half of the Ss did the lesson, then the writing exercises, then the test; the other half did the lesson and then took the test. Appendix C is a writing exercise to accompany Lesson 16.

It is important to note that throughout the project, Ss given a particular experimental treatment were divided evenly between the treatment groups of a subsequent experiment.

Table 2 presents the sequence of events which took place during the project.

DESIGN OF EXPERIMENTS AND DISTRIBUTION OF Ss TO TREATMENTS

It has been stated previously that the data from 48 Ss were used for purposes of statistical analysis of the experiments. These Ss were eight boys and eight girls from each of three IQ ranges. Sex and IQ were used as stratifying variables in each of the four experiments.

Experiment 1

The addition of the advance organizer variable completed the design of Experiment 1 which was thus a $3 \times 2 \times 2$ factorial design. The factors were high, medium, and low IQ; sex; and advance organizers and control material. After Pretest I, the eight Ss within a sex-IQ group were ranked on the basis of their pretest scores. The top two scores in the group were randomly split between advance organizers and control material. Then the next two highest scorers in the group were randomly split. This procedure was followed until each of the six sex-IQ groups had been split between the advance organizer treatment and the control treatment.

Experiment 2

Lesson 10 was made up of 25 frames which presented new material followed by 25 review frames. Experiment 2 consisted of giving half of the Ss Lesson 10 in its entirety before Posttest I while the other half of the Ss were given the posttest immediately after they had completed the first 25 frames of Lesson 10. Because of the $3 \times 2 \times 2$ design of Experiment 1, there were 12 groups with four Ss in a group and each group was homogeneous with respect to sex, IQ, and advance organizers. Each of these groups was randomly divided in half; two of the Ss were assigned to the review treatment and two were assigned to the no-review treatment. Thus, Experiment 2 was a $3 \times 2 \times 2 \times 2$ factorial design with IQ, sex, advance organizers and review as the four factors.

After Posttest I, the Ss in the no-review treatment were given the review frames from Lesson 10. It was assumed that, in terms of review information, this procedure would equate these Ss with those who had had the review before Posttest I.

Experiment 3

Experiment 3 tested the effects of the addition of negative instances to Lessons 11-15 and feedback of test results. Because all Ss following Experiment 2 were eventually given the review frames, it was assumed that any differences produced by the review information were again balanced at the beginning of Experiment 3. (A later analysis showed a nonsignificant effect of review.) Thus, there were still 12 groups with four Ss in a group, identifiable by sex and IQ. Each of the four Ss in these 12 groups was randomly assigned to one of the four treatments in Experiment 3: negative instances and feedback, negative instances and no feedback, regular instances and feedback, or regular instances and no feedback. Experiment 3 was thus a $3 \times 2 \times 2 \times 2$ factorial design with IQ, sex, negative instance treatment, and feedback treatment as the four factors.

Experiment 4

Experiment 4 tested the effect of daily exercises relevant to Lessons 16-21. In order to assign Ss to treatments, Ss within a sex-IQ group were ranked on the basis of their scores on Posttest II which was given after Lesson 15. The top two Ss in a group were divided; one was assigned randomly to the exercise group

Table 2

Sequence of Events

	Date	Event
Experiment 1	June 20	Pretest I covering Lessons 1-10
	June 21	1/2 Ss given advance organizer, while the other 1/2 read filler material; all Ss do Lesson 1.
	June 22	Advance organizer or filler, Lesson 2, Test 2.
	June 23	Advance organizer or filler, Lesson 3, Test 3.
	June 24	Advance organizer or filler, Lesson 4, Test 4.
	June 27	Advance organizer or filler, Lesson 5, Test 5.
	June 28	Advance organizer or filler, Lesson 6, Test 6.
	June 29	Advance organizer or filler, Lesson 7, Test 7.
	June 30	Advance organizer or filler, Lesson 8, Test 8.
	July 1	Advance organizer or filler, Lesson 9, Test 9.
Experiment 2	July 5	Advance organizer or filler, Lesson 10; 1/2 Ss given review of Lessons 1-10 and then Posttest I, covering Lessons 1-10; the other 1/2 of the Ss given Posttest I immediately after completing Lesson 10, and then the review.
Experiment 3	July 6	Pretest II covering Lessons 11-21.
	July 7	1/2 Ss given negative instances in Lesson 11, the other 1/2 given the regular program; all Ss given Test 11.
	July 8	Lesson 12: 1/4 Ss given negative instances in Lesson 12 and feedback of test results on Test 11; 1/4 Ss negative instances and no feedback of test results; 1/4 Ss regular program and feedback; 1/4 Ss regular program and no feedback. All Ss given Test 12.
	July 11	Lesson 13: Negative instances or regular program, feedback on Test 12 or no feedback; Test 13.
	July 12	Lesson 14: Negative instances or regular program, feedback on Test 13 or no feedback; Test 14.
	July 13	Lesson 15: Negative instances or regular program, feedback on Test 14 or no feedback; Test 15; Posttest II covering Lessons 11-15.
Experiment 4	July 14	Lesson 16: 1/2 Ss given exercises relevant to Lesson 16, then Test 16; the other 1/2 Ss do Test 16 directly after Lesson 16.
	July 15	Lesson 17: exercise on Lesson 17, or no exercise, Test 17.
	July 18	Lesson 18: exercise on Lesson 18, or no exercise, Test 18.

July 19	Lesson 19: exercise on Lesson 19, or no exercise, Test 19.
July 20	Lesson 20: exercise on Lesson 20, or no exercise, Test 20.
July 21	Lesson 21: exercise on Lesson 21, or no exercise, Test 21.
July 22	All Ss given general writing exercise, Posttest IIB, covering Lessons 16-21, and questionnaire on Project.

and the other to the no-exercise group. This procedure was followed with each subsequent pair of Ss in the sex-IQ group and with all of the other sex-IQ groups until all Ss had been assigned randomly to exercises or no exercises. Experiment 4 was a 3 × 2 × 2 factorial design with IQ, sex, and exercises as the three factors.

RESULTS

INTRODUCTION

The results section is organized under four headings. Program Measures gives the data regarding mean time and mean error rate for each of the 21 lessons. Test Score Results contains the analyses of the pretests, the posttests, the 19 daily tests, and the exercise test. Item Analysis gives a short summary of the item analyses done on the tests. The fourth section, Questionnaire Responses, is an examination of a questionnaire given to the Ss upon completion of the project, designed to assess Ss' attitudes toward the project. Interpretation of the results will be presented in Chapter IV, Discussion and Summary.

PROGRAM MEASURES

One of the primary purposes for the Summer English Project was to collect descriptive data which would provide information for revision of the programed lessons. The two descriptive measures taken for the lessons were (1) time per lesson and (2) errors per lesson. The usefulness of the program in a classroom situation is in part dependent on the amount of time required to work through the lessons. A lesson demanding more classroom time than is normally available, as well as a lesson which takes an inordinately short period of time, would seem to require revision. From the standpoint of motivation and interest, an exorbitantly high error rate on responses to frames would also be undesirable. Given the range of mental ability which one is likely to find in a classroom, each lesson should be constructed so that within one class period it can be completed by every student working at a steady pace and yet will present a fairly substantial task.

To provide information on the average amount of time required to work steadily through each lesson, Ss were asked to record the exact times at which they began and completed the day's lesson. Table 3 presents the mean time in minutes for males and females in the three ability groups. The overall mean time for the high group was 20.3; the mean time for the middle, 21.7; and the mean time for the low, 23.4. The times given in Table 3 are, of course, a function of the number of frames contained in the lessons, which varied in length from 41 frames to 85. Shown in Table 4 is the mean time for the lessons grouped according to five ranges of length. As can be seen, the amount of time spent on shorter lessons was approximately 16 minutes, while the time spent on the longest lessons was approximately 30 minutes.

To provide a fairly accurate means of assessing the error rates for each lesson, Ss recorded their responses to the frames in ink on a separate answer sheet. Rather than erasing incorrect responses, Ss were instructed to cross them out before writing the correct response. For each lesson, the mean error rates for sex and ability groups were calculated from a tally of the number of incorrect responses. A table giving the exact number of errors per lesson appears in Appendix D.

The mean error rates presented in Table 5 show that the high group had an average of 4.6 per cent incorrect responses per lesson; the middle, 6.4 per cent; and the low, 10.9 per cent. The overall error rate, then, was 7.3 per cent.

TEST SCORE RESULTS

Analyses of variance (ANOVA) were used to test the statistical significance of the main effects and interactions in each experimental design. One-hundred fourteen ANOVA were computed from the test scores using the BMD 14 General Linear Hypothesis Computer Program at the University of Wisconsin Computing Center.

Experiment I

Ss' scores on Pretest I, Daily Tests 2 through

Table 3

Mean Time Per Lesson

Ss in Group	Group	Lesson 1	2	3	4	5	6	7	8
		# Frames 60	50	65	49	63	85	53	68
9	Hi Male	30.3	21.0	24.3	15.2	23.3	29.3	15.6	22.7
11	Hi Female	27.8	20.8	24.0	15.5	24.4	30.2	17.8	22.0
20	Total Hi	29.1	20.9	24.1	15.3	23.9	29.7	16.7	22.3
9	Mid Male	34.2	22.7	28.6	18.1	25.6	32.4	20.3	24.3
10	Mid Female	30.1	22.9	26.9	16.9	25.5	31.6	20.5	24.0
19	Total Mid	32.1	22.8	27.7	17.5	25.5	32.0	20.4	24.2
10	Lo Male	37.4	26.6	31.3	20.5	29.7	35.9	20.8	25.4
10	Lo Female	33.1	22.9	26.8	19.0	25.1	32.6	19.4	22.2
20	Total Lo	35.2	24.7	29.0	19.7	27.4	34.2	20.1	23.3

Lesson	9	10	11	12	13	14	15	16	17	18	19	20	21
# Frames	63	58	61	60	47	41	54	57	62	81	64	46	46
	17.7	18.9	23.7	22.7	14.8	12.9	20.2	19.6	19.7	26.2	20.0	12.9	12.6
	18.9	22.1	23.7	24.0	18.7	16.2	21.2	18.9	19.2	25.7	20.6	12.0	12.7
	18.3	20.5	23.7	23.3	16.7	14.5	20.7	15.2	19.4	25.9	20.3	12.4	12.6
	20.6	24.8	25.6	24.2	16.9	15.9	21.3	20.9	22.4	26.6	19.4	13.4	14.7
	19.6	22.2	26.3	25.9	18.4	18.4	22.6	22.8	21.3	27.5	21.3	14.0	14.3
	20.1	28.5	25.9	25.0	17.6	17.1	21.9	21.8	21.8	27.0	20.3	13.7	14.5
	20.7	27.1	25.7	28.9	19.2	17.9	25.4	24.7	24.7	27.9	19.9	16.3	17.2
	18.2	24.1	23.3	21.7	16.1	17.2	19.9	22.2	20.6	25.6	20.3	15.7	15.9
	19.4	25.6	24.5	25.3	17.6	17.5	22.6	23.4	22.6	26.7	20.1	16.0	16.5

Table 4

Mean Time for Lessons of Varying Lengths for Three Ability Groups

Group	# Frames per Lesson	40-49	50-59	60-69	70-79	80-89
	# Lessons	5	5	9	0	2
High		14.3	18.8	22.7	-	27.8
Middle		16.1	23.1	24.7	-	29.5
Low		17.5	23.3	25.2	-	30.5
Combined		15.9	21.7	24.2	-	29.2

9, and Posttest I were subjected to ANOVA. Three ANOVA were computed from scores on each test: one analysis for the multiple choice (MC) part, one for the completion (Comp) part, and one for the Total of the two parts (MC + Comp). The sum of scores on the Daily Tests was treated as a single score and analyzed as was this sum plus the score of Posttest I. A total of 36 ANOVA were relevant to Experiment 1.

IQ. Experiment 1 was a 3 × 2 × 2 factorial design; the factors were IQ, Sex, and Advance Organizer (AO). Pretest I was analyzed to see if any differences existed between and among experimental groups. The IQ factor was significant at the .01 level. On Pretest I, as

Table 5

Mean Error Rate[a]

Ss/Grp	Group	Lesson	1	2	3	4	5	6	7	8	9	10	11
		# Frames	60	50	65	49	63	85	53	68	63	58	61
9	Hi Male		.00	.05	.06	.03	.04	.04	.02	.05	.05	.06	.06
11	Hi Female		.00	.05	.07	.02	.04	.06	.04	.05	.05	.06	.06
20	Total Hi		.00	.05	.07	.02	.04	.05	.03	.05	.05	.06	.06
9	Mid Male		.06	.09	.10	.06	.05	.06	.04	.05	.05	.09	.06
10	Mid Female		.02	.07	.09	.03	.05	.08	.04	.06	.07	.08	.07
19	Total Mid		.04	.08	.10	.04	.05	.07	.04	.06	.06	.08	.07
10	Lo Male		.13	.15	.20	.10	.14	.13	.08	.11	.11	.15	.15
10	Lo Female		.10	.11	.14	.08	.10	.09	.08	.10	.08	.10	.06
20	Total Lo		.11	.13	.17	.09	.12	.11	.08	.10	.09	.13	.11
59	Grand Total		.05	.09	.11	.05	.07	.08	.05	.07	.07	.09	.08

Lesson	12	13	14	15	16	17	18	19	20	21
# Frames	60	47	41	54	57	62	81	64	46	46
	.07	.09	.08	.05	.03	.05	.05	.05	.05	.04
	.07	.09	.05	.05	.03	.04	.04	.05	.03	.03
	.07	.09	.06	.05	.03	.04	.04	.05	.04	.03
	.09	.10	.07	.07	.07	.04	.05	.04	.04	.04
	.10	.10	.11	.08	.05	.03	.06	.05	.04	.05
	.10	.10	.09	.08	.06	.04	.06	.05	.04	.05
	.18	.20	.14	.14	.11	.10	.11	.11	.09	.12
	.09	.09	.12	.07	.04	.05	.07	.06	.05	.07
	.14	.14	.13	.10	.08	.08	.09	.08	.07	.09
	.10	.11	.10	.08	.06	.05	.06	.06	.05	.06

[a] $M.E.R. = \frac{\#\ \text{Errors}}{\#\ \text{In Grp} \times \#\ \text{Frames}}$

shown in Table 6, the High, Medium, and Low IQ groups averaged 47.19, 32.75, and 28.06 points, respectively, out of 75 possible points. There were no significant F ratios for the sex factor, the AO factor, or for any of the interactions on Pretest I. Table 7 summarizes the means, standard deviations, and F ratios for each of the analyses of Pretest I.

On the average, Ss in the high IQ range scored 4.22 more correct answers on Tests 2-9 than did Ss in the medium IQ range. The latter Ss scored, on the average, 2.76 more correct answers than Ss in the low IQ range. The means for the three groups were 19.02, 14.80, and 12.04, respectively, where the highest possible score on each of Tests 2-9 was 24 correct answers. Table 6 presents the means and standard deviations for the High, Medium, and Low IQ groups; for males and females; and for the AO and control groups. These measures are given for each of the Daily Test Total Scores and for the Pretest and Posttest Total Scores. The means for High, Medium, and Low IQ groups are graphed in Figure 1.

Sex There was a trend for the sex factor to be significant on Tests 2-9. In the analyses, sex was significant beyond the .10 level 12 times; in 4 of these 12 analyses, the sex factor was significant at the .05 level. The 4 analyses shown in Table 7 were those on Total Score of Test 2, MC of Test 4, MC of Test 5, and the sum of the Total Scores for Tests 2-9. In every case, females were slightly above

Table 6

Means and Standard Deviations by Groups
for Tests 1 through 9, Pretest I, and Posttest I

		High	Med	Low	Male	Female	A.O.	No. A.O.
Pretest I	Mean	47.19	32.75	28.06	36.29	35.71	36.58	35.42
	St. Dev.	7.8	6.6	4.9	11.0	10.0	10.6	10.4
Posttest I	Mean	58.50	44.25	35.31	44.04	48.00	46.96	45.08
	St. Dev.	6.9	10.5	7.2	13.8	11.3	12.5	13.0
Test 1 & 2	Mean	20.44	16.31	13.81	15.92	17.79	17.38	16.33
	St. Dev.	2.2	3.8	3.3	4.7	3.3	4.1	4.3
Test 3	Mean	19.12	15.06	11.06	14.67	15.50	14.96	15.21
	St. Dev.	3.2	2.8	4.2	5.0	4.6	4.5	5.1
Test 4	Mean	19.94	16.06	13.56	15.50	17.54	16.17	16.88
	St. Dev.	2.7	3.9	4.0	4.6	4.1	4.2	4.6
Test 5	Mean	20.06	13.88	11.31	14.08	16.08	14.62	15.54
	St. Dev.	1.8	5.1	3.3	5.5	4.6	5.3	5.1
Test 6	Mean	19.06	15.75	13.50	15.71	16.50	16.00	16.21
	St. Dev.	2.5	3.6	2.9	3.5	4.0	3.5	4.1
Test 7	Mean	18.25	15.06	11.25	14.62	15.08	15.29	14.42
	St. Dev.	1.9	3.2	2.5	4.1	3.7	3.4	4.3
Test 8	Mean	19.88	16.06	12.69	15.46	16.96	16.33	16.08
	St. Dev.	2.5	3.2	4.0	4.6	4.1	4.7	4.1
Test 9	Mean	15.44	10.19	9.12	11.62	11.54	11.83	11.33
	St. Dev.	2.5	2.7	3.1	3.5	4.3	4.2	3.7
Σ Tests 1-9	Mean	152.19	118.38	96.31	117.58	127.00	122.58	122.00
	St. Dev.	12.3	22.3	20.5	31.1	28.1	28.8	31.2

males in number of correct answers (see Figure 2). Females, on the average, scored 1.18 points higher than males on Tests 2-9.

Advance Organizer None of the analyses showed a significant difference between the AO treatment and the control treatment. The Ss who received the AO consistently scored about the same as those who received the control material. These results are indicated in Figure 3 on a day-by-day basis.

In all the analyses, only one interaction reached the .05 level of significance. This was a Sex × IQ interaction on the Comp part of Test 7. The relevant means were:

Males High 10.12 Med 7.12 Low 5.37
Females High 9.12 Med 9.25 Low 5.25

On Posttest I, IQ was significant at the .01 level as would be expected. No other main effects or interactions reached the .05 level of significance.

Pretest I compared to Posttest I Figure 4, Figure 5, and Figure 6 indicate the increase in Total Scores from Pretest I to Posttest I. The increases can also be read from Table 6 by comparing the Pretest score with the Posttest. Note that females improved more than males, and that the Low group advanced comparatively less than the High and Medium. The improvement of the AO group was similar to that of the

Table 7

Summary Table of Means, Standard Deviations, and F Ratios
for Pretest, Daily Tests, and Posttests
(Lessons 1-10)

				F Ratios						
		Mean	St. Dev.	IQ	Sex	AO	IQ × Sex	IQ × AO	Sex × AO	IQ × Sex × AO
Pretest I	MC	23.23	6.33	30.39**						
	Comp	12.77	4.87	18.39**					3.86	
	Total	36.00	10.42	31.79**						
Tests 1 & 2	MC	9.08	1.93	8.73**	3.68					
	Comp	7.77	2.53	20.19**	3.06					
	Total	16.85	4.16	17.87**	4.21*					
Test 3	MC	7.21	2.23	8.76**						
	Comp	7.88	2.83	26.73**						
	Total	15.08	4.76	19.90**						
Test 4	MC	8.94	2.27	10.17**	5.37*					
	Comp	7.58	2.60	7.92**						
	Total	16.52	4.40	11.99**	3.63					
Test 5	MC	8.08	2.79	17.80**	7.01*	3.28				
	Comp	7.00	2.71	20.53**						
	Total	15.08	5.16	23.55**	3.49					
Test 6	MC	8.31	1.82	10.89**						
	Comp	7.79	2.41	8.35**						
	Total	16.10	3.77	13.01**						
Test 7	MC	7.15	1.87	11.81**						
	Comp	7.71	2.53	24.36**			3.28*			
	Total	14.85	3.84	32.05**			3.22			
Test 8	MC	7.25	2.62	9.50**	3.58					
	Comp	8.96	2.32	15.83**					2.93	
	Total	16.21	4.37	18.28**					2.94	
Test 9	MC	5.42	2.02	9.46**				2.76		
	Comp	6.17	2.33	21.60**						
	Total	11.58	3.89	21.60**						
Posttest I	MC	28.44	8.09	25.49**						
	Comp	17.58	5.09	22.87**	3.53					
	Total	46.02	12.65	28.33**						
Σ Tests 2-9	MC	61.44	13.66	23.90**	4.54*					
	Comp	60.85	16.68	35.15**						
	Total	122.29	29.68	32.36**						
Σ Tests 2-9 + Posttest	MC	89.88	21.06	28.31**	3.63					
	Comp	78.44	21.35	34.81**						
	Total	168.31	41.60	34.08**	2.88					

*p < .05
**p < .01

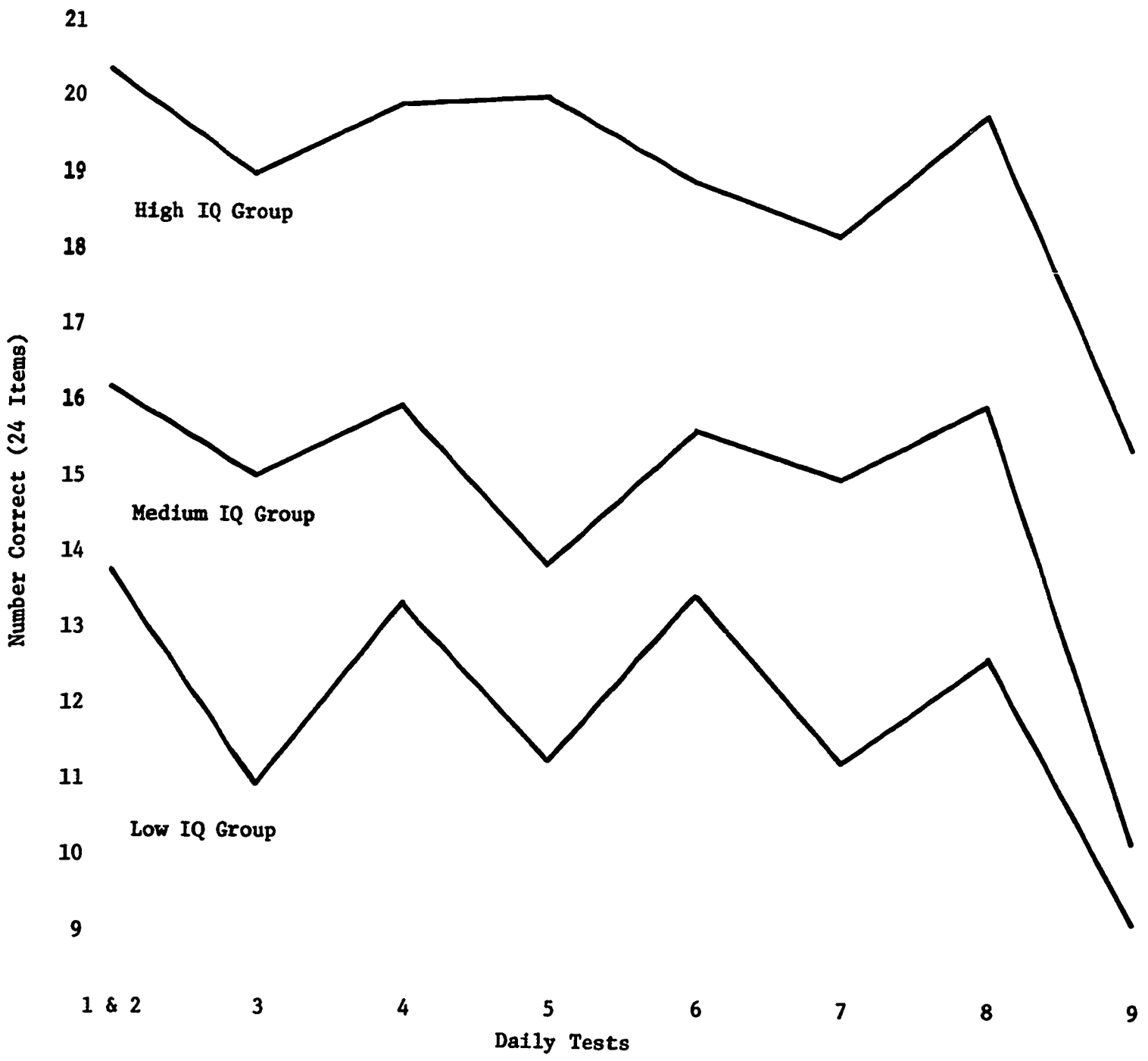

Fig. 1. The average Total Scores for High, Med, and Low IQ groups on Daily Tests 1 through 9.

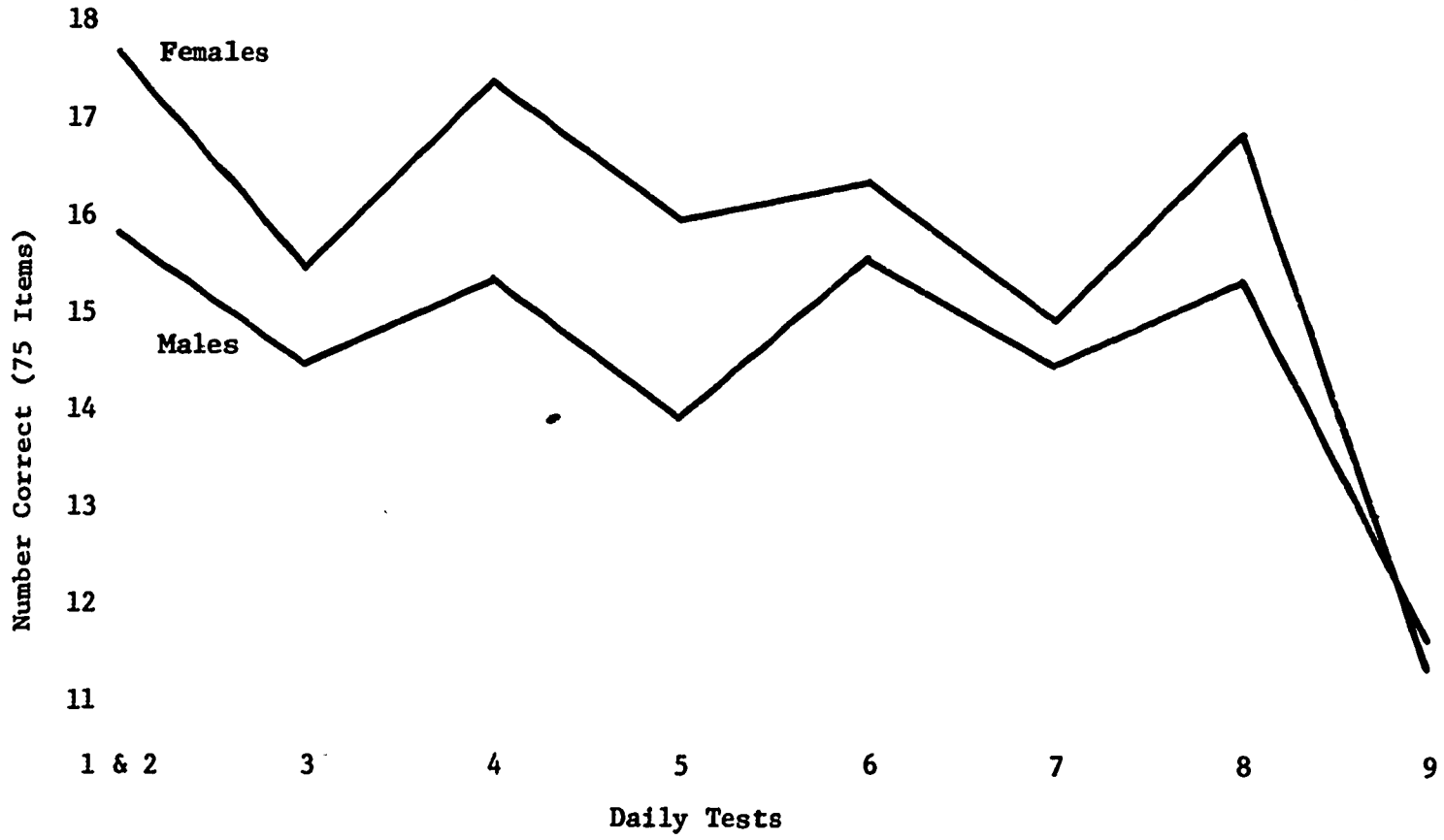

Fig. 2. The average Total Scores for Males and Females on Daily Tests 1 through 9.

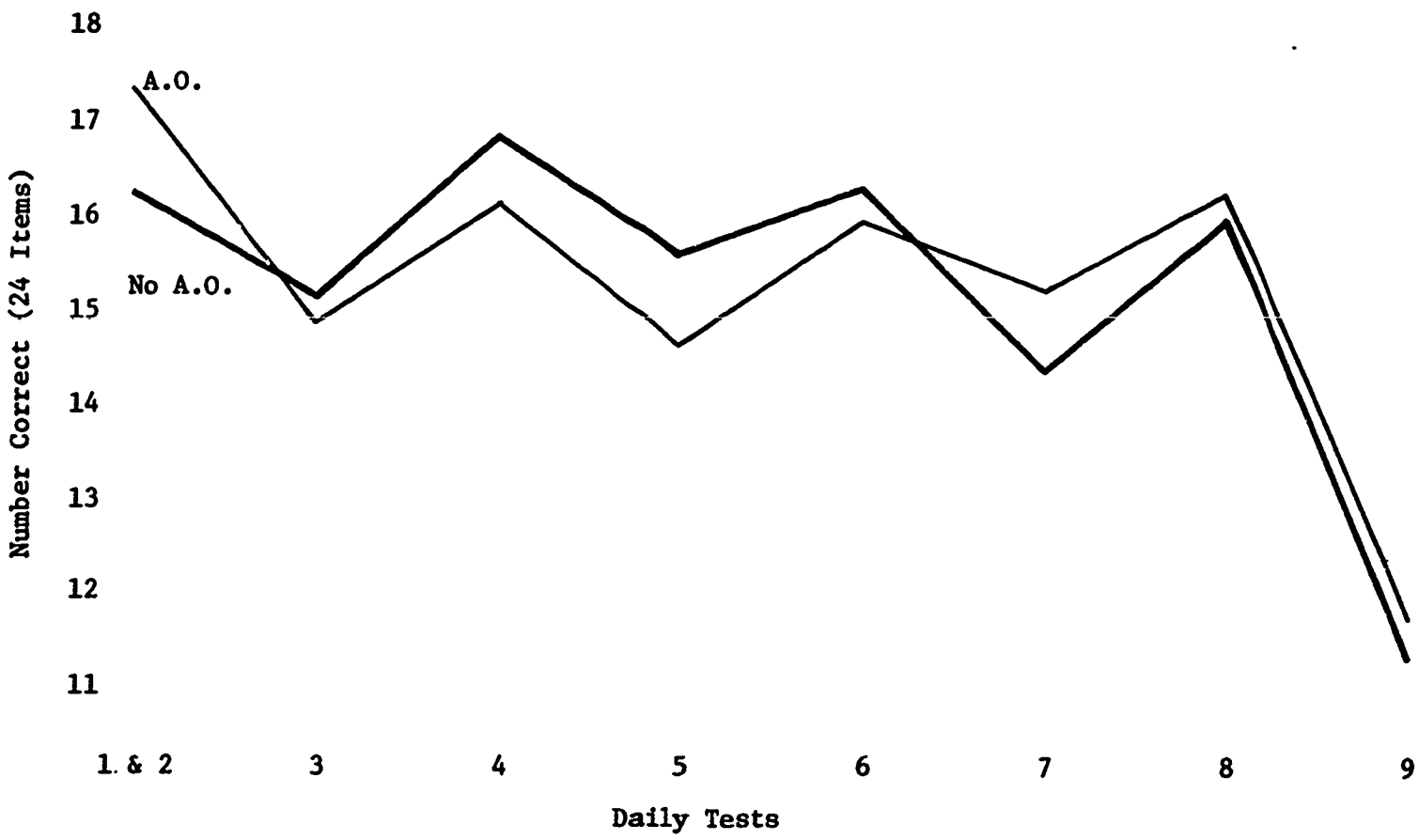

Fig. 3. The average Total Scores for Advance Organizer (A. O.) and No A. O. groups on Daily Tests 1 through 9.

No AO. The average increase in Total Score for all Ss was 10 points (from 36.0 to 46.0 correct out of a possible 75).

Experiment 2

Posttest I was analyzed a second time. On this ANOVA, Review—No Review (Rev) was included as a factor making a 3 × 2 × 2 × 2 factorial analysis. The means for Total Score were as follows: High 57.56; Medium 45.50; Low 35.06; Males 43.50; Females 48.58; AO 47.50; No AO 44.58; Rev 46.33; No Rev 45.75. IQ was significant at the .01 level; sex approached significance; and AO and Rev were not significant. The interaction of IQ × Sex × AO × Rev was significant at the .05 level.

Experiment 3

Thirty-five ANOVA were relevant to Experiment 3. This sum includes three analyses each of Pretest IIA and IIB, Pretest IIA, Posttest IIA, Daily Tests 11 through 16, the sum of the Daily Tests, and sum of Daily Tests plus Posttest IIA. One ANOVA was also performed on the Recognition Test given after Lesson 17, and one on the sum of Lesson 16 plus Recognition Test plus Posttest IIA. Many of the factors and interactions were significant. The rest of this devoted to presenting these data. Reference to Tables 8 and 9 will prove helpful. Experiment 3 involved a 3 × 2 × 2 × 2 design which included the factors of IQ, Sex, Positive Instances only (No Neg) or one-fourth Negative (Neg), and Feedback (Fdbk) or No Fdbk.

IQ. On Pretest IIA and IIB, IQ was the only significant factor. The High, Medium and Low groups scored 46.00, 35.75, and 25.31 correct, respectively, out of a possible 90 points. The other means and standard deviations are given in Table 8. No other main effect or interaction reached the .05 level of significance on Pretest IIA and B.

IQ was significant beyond the .01 level in every analysis (except for MC of Test 12, where it reached .05 but not .01); the reader is referred to Table 9 for these F ratios and to Table 8 for the means and standard deviations of the Total Score for each of the main effects. Figure 7 is a graph of the Total Score of each IQ group on Daily Tests 11–16.

Test II On MC of Test 11, in addition to IQ, sex was significant at the .01 level. Males scored an average of 5.88 correct, and females 7.08 (see Table 8 and Figure 8). The interaction of IQ × Fdbk was also highly significant on this subtest. The means were as follows:

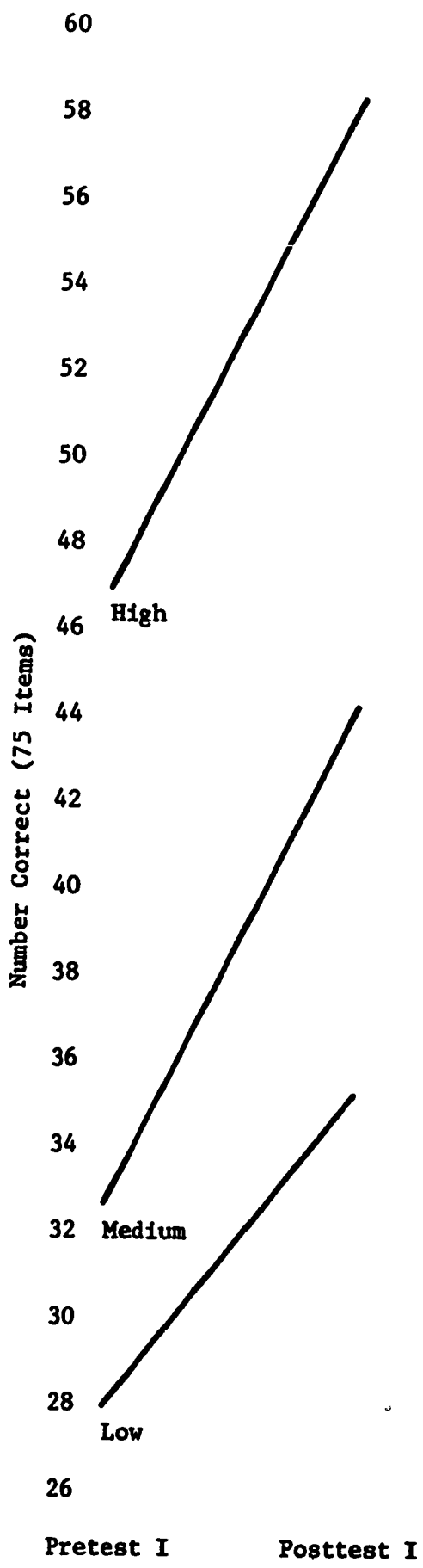

Fig. 4. The average Pretest I and Posttest I Total Scores for High, Med, and Low IQ groups.

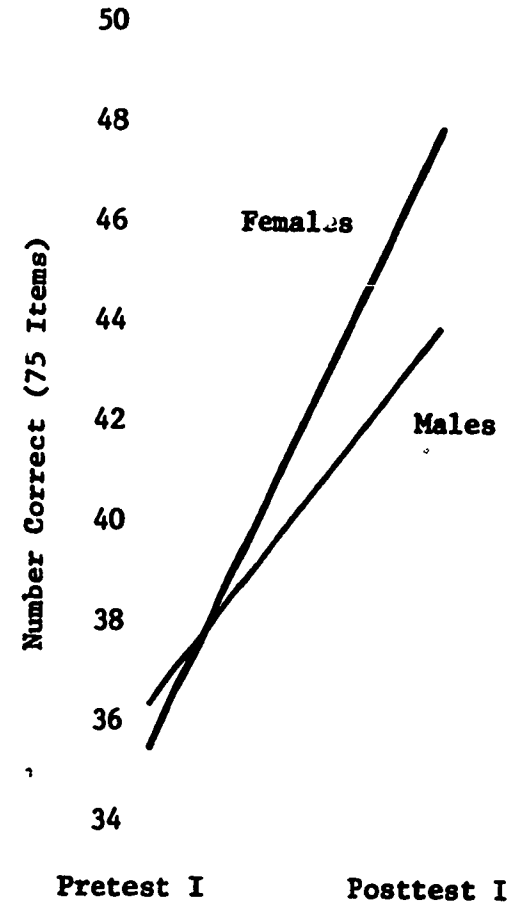

Fig. 5. The average Pretest I and Posttest I Total Scores for Males and Females.

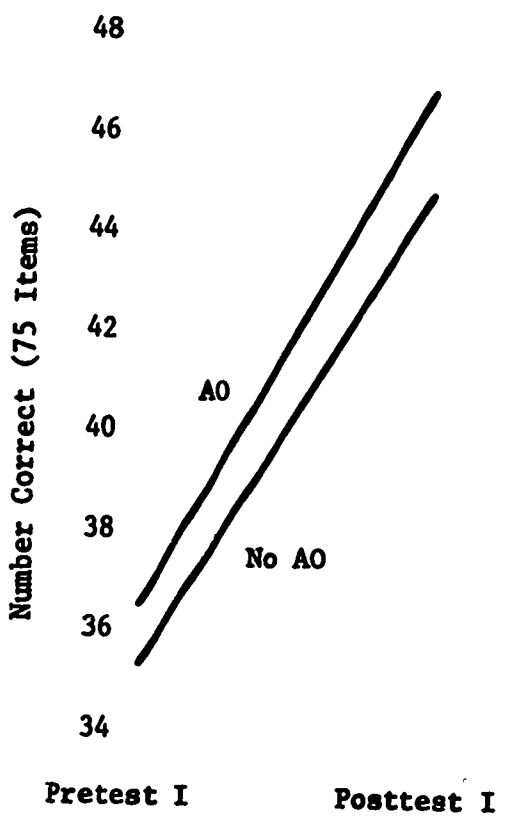

Fig. 6. The average Pretest I and Posttest I Total Scores for Advance Organizer (AO) and No AO groups.

Table 8

Means and Standard Deviations by Groups for Tests 11 through 17, Pretest IIA & B, Pretest IIA and Posttest IIA

		High	Med	Low	Male	Female	Fdbk	No Fdbk	Neg	No Neg
Pretest IIA & B										
	Mean	46.00	35.75	25.31	35.58	35.79	34.83	36.54	35.38	36.00
	St. Dev.	8.4	8.6	5.9	11.0	12.0	12.6	10.2	10.6	12.4
Prestest IIA										
	Mean	21.56	16.31	12.38	16.92	16.58	16.46	17.04	16.79	16.71
	St. Dev.	3.9	3.1	3.9	4.7	5.8	5.9	4.6	4.9	5.6
Posttest IIA										
	Mean	28.25	22.69	16.94	21.46	23.79	23.71	21.54	22.21	23.04
	St. Dev.	5.8	4.2	4.0	7.3	5.7	7.5	5.4	7.1	6.1
Test 11										
	Mean	17.94	12.38	9.88	12.71	14.08	13.38	13.42	13.00	13.79
	St. Dev.	3.1	3.7	3.4	5.2	4.3	5.6	3.8	5.3	4.1
Test 12										
	Mean	16.62	11.88	10.06	12.04	13.67	13.17	12.54	12.17	13.54
	St. Dev.	5.2	5.0	3.2	6.0	4.3	5.8	4.8	5.6	4.9
Test 13										
	Mean	16.88	10.69	9.69	11.71	13.12	12.12	12.71	11.75	13.08
	St. Dev.	3.2	3.8	2.3	4.6	4.3	5.4	3.4	4.9	4.1
Test 14										
	Mean	17.06	11.31	9.50	11.21	14.04	13.04	12.21	12.17	13.08
	St. Dev.	4.2	5.0	3.5	5.6	4.7	6.1	4.5	5.8	4.9
Test 15										
	Mean	17.44	11.50	8.31	11.75	13.08	13.29	11.54	12.12	12.71
	St. Dev.	4.0	3.7	3.3	5.5	5.0	5.8	4.6	6.1	4.4
Test 16										
	Mean	19.50	13.75	10.31	13.12	15.92	14.83	14.21	14.71	14.33
	St. Dev.	3.6	3.2	3.5	5.2	4.7	5.5	4.7	5.7	4.5
Recognition Test 17										
	Mean	12.88	6.31	5.25	8.75	7.54	8.08	8.21	7.71	8.58
	St. Dev.	5.8	4.0	3.5	5.9	5.3	6.0	5.3	5.6	5.7
Σ (Total 11-16)										
	Mean	105.44	71.50	57.75	72.54	83.92	79.83	76.62	75.92	80.54
	St. Dev.	20.4	21.5	15.8	30.1	24.5	32.2	23.0	31.3	24.2
Σ (11-16 & Posttest IIA)	Mean	133.69	94.19	74.69	94.00	107.71	103.54	98.17	[illegible]	103.58
	St. Dev.	24.9	24.9	18.5	36.2	29.6	38.9	27.6	[illegible]	29.5
Test 16 & Recog & Posttest IIA										
	Mean	60.62	42.75	32.50	43.33	47.25	46.62	43.96	44.62	45.96
	St. Dev.	12.2	9.2	8.5	16.9	13.7	17.4	13.3	16.3	14.7

Fdbk	High 9.25	Med 4.75	Low 5.50
No Fdbk	High 7.75	Med 7.25	Low 4.38

No other main effect or interaction was significant (See Table 8, Figure 9, and Figure 10). On Comp of Test 11, sex was not significant and neither was the IQ × Fdbk interaction. Two 3-way interactions were, however, significant at the .05 level. One of them, the IQ × Sex × Neg interaction, had been nearly significant on Pretest IIA and B. This interaction could be explained as follows. Medium and Low males did poorer when they received Neg than when they received No Neg. High and Medium females, on the other hand, performed a little better with the Neg than without.

The second 3-way interaction significant at the .05 level was IQ × Fdbk × Neg. This resulted from some erroneous split of Ss since no group had as yet received any Fdbk.

Both of these complex interactions remained significant when the Total Score for Test 11 was

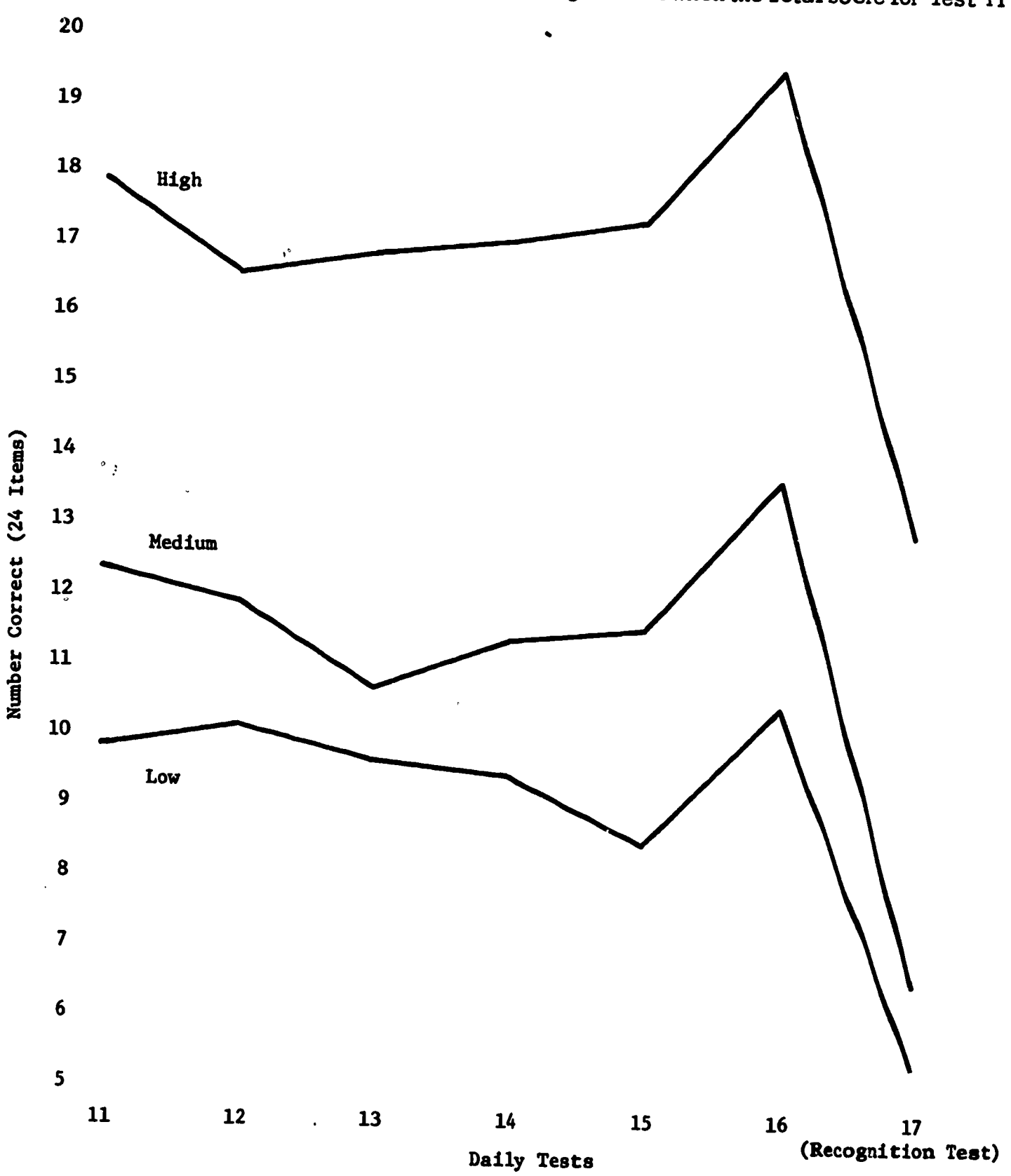

Fig. 7. The average Total Scores for High, Med, and Low IQ groups on Daily Tests 11 through 16 and Recognition Test 17.

Table

Summary Table of Means, Standard
for Pretest, Daily
(Lessons

				F			
		Mean	St. Dev.	IQ	Sex	Fdbk	Neg
Pretest IIA & B	MC	22.85	6.76	16.25**			
	Comp	12.83	5.60	20.48**			
	Total	35.69	11.39	27.60**			
Test 11	MC	6.48	2.38	26.48**	8.67**	3.72	
	Comp	6.92	2.83	20.91**			
	Total	13.40	4.75	33.63**			
Test 12	MC	6.71	2.43	5.31*	3.63		
	Comp	6.15	3.33	13.26**			
	Total	12.85	5.26	11.59**			
Test 13	MC	6.58	2.32	20.21**			
	Comp	5.83	2.60	31.72**	6.03*		
	Total	12.42	4.48	35.06**	3.48		3.08
Test 14	MC	6.06	2.53	9.84**	6.35*		
	Comp	6.56	3.18	18.57**	4.52*		
	Total	12.62	5.31	17.92**	6.92*		
Test 15	MC	5.85	2.46	39.49**		6.72*	
	Comp	6.58	3.21	19.55**	4.82*		
	Total	12.42	5.23	38.13**		4.08	
Pretest IIA	MC	11.52	3.73	9.16**			
	Comp	5.19	2.47	18.10**			
	Total	16.75	5.22	22.17**			
Posttest IIA	MC	14.79	4.48	17.91**	3.17		
	Comp	7.79	2.70	47.12**	3.70	8.33**	
	Total	22.62	6.57	34.51**	4.40*	3.80	
Test 16	MC	6.71	2.90	24.34**	3.63		
	Comp	7.81	2.69	29.47**	13.03**		
	Total	14.52	5.09	38.76**	10.51**		
Test 17 Recognition Test		8.15	5.60	15.25**			
Total (Test 16 & Recog. & Posttest IIA)		45.29	15.34	43.29**			
Σ (Tests 11-16)	MC	38.40	12.38	36.56**	5.91*		
	Comp	39.85	15.93	35.64**	5.59*		
	Total	78.25	27.75	39.07**	6.29*		
Tests 11-16 & Posttest IIA							
	MC	53.19	16.12	37.12**	6.16*		
	Comp	47.65	18.05	41.54**	6.06*		
	Total	100.85	33.47	43.12**	6.73*		

*p < .05

**p < .01

Deviations and F Ratios
Tests, and Posttests
11-15)

Ratios										
IQ × Sex	IQ × Fdbk	IQ × Neg	Sex × Fdbk	Sex × Neg	Fdbk × Neg	IQ×Sex × Fdbk	IQ×Sex × Neg	IQ×Fdbk × Neg	Sex× Fdbk × Neg	IQ×Sex × Fdbk × Neg
							3.21			
							3.09			
	9.66**	3.35					3.10			
			3.48				3.50*	4.22*		
	5.45*						4.37*	3.43*		
	6.26**			3.10						
	3.77*		5.12*			4.45*	3.11			
	5.70**		4.27*			2.68	2.57			
			9.18**	5.88*	3.30	6.64**	2.97	5.23*		
										2.82
	3.09		5.83*			3.40*	2.68			
						2.65				
	3.14					3.30	3.30			
	3.13					3.64*	2.93			
	3.88*				4.74*		2.55			
	2.75						3.15			
	3.88*						3.64*			
							3.25			
	9.20**									
						3.02				
	8.92**									
	3.35			3.63						
	3.98*	2.57	3.42							
	4.77*						2.68			
	3.15						2.78			3.06
	7.93**						3.37			
	7.82**						3.46*			
	4.52*		4.22			2.85	4.48*			
	6.26**					2.80	4.32*			
	9.76**					2.61	2.81			
	4.70*		3.58			3.11	4.46*			
	7.58**					3.07	4.04*			

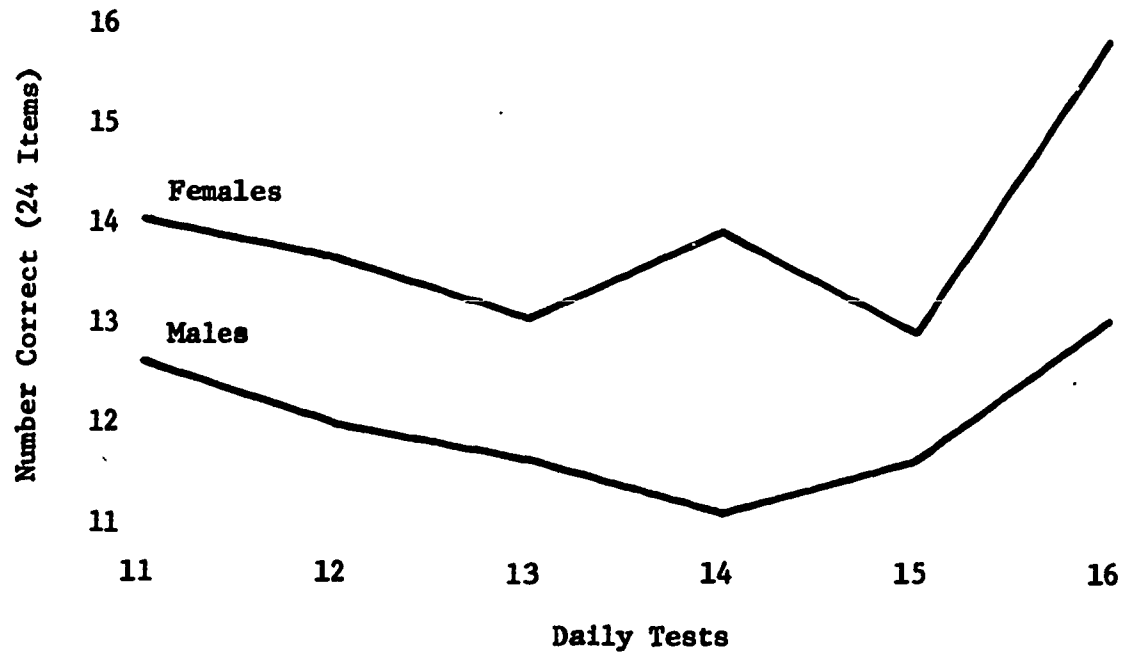

Fig. 8. The average Total Scores for Males and Females on Daily Tests 11 through 16.

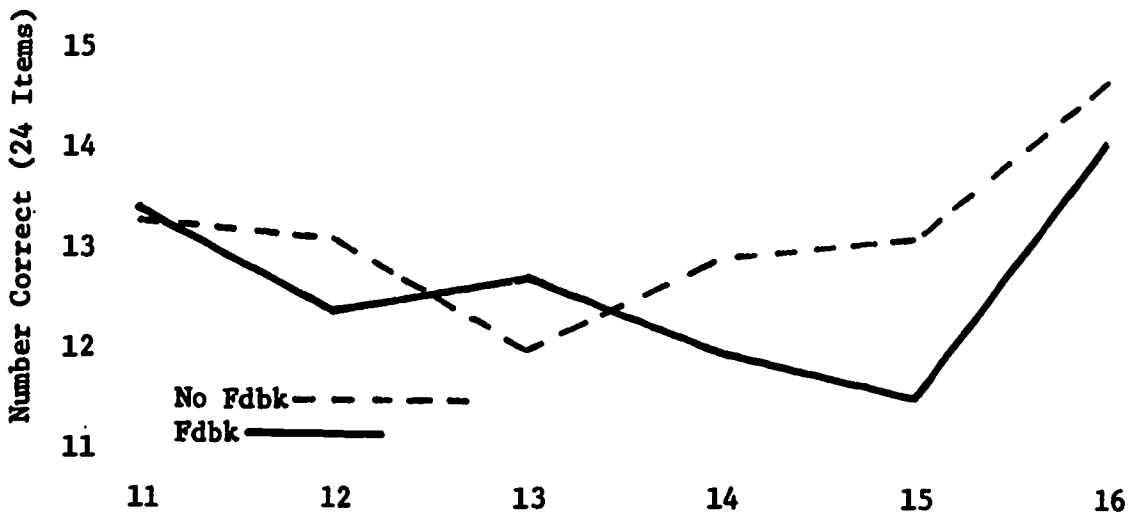

Fig. 9. The average Total Scores for Feedback (Fdbk) and No Fdbk groups on Daily Tests 11 through 16.

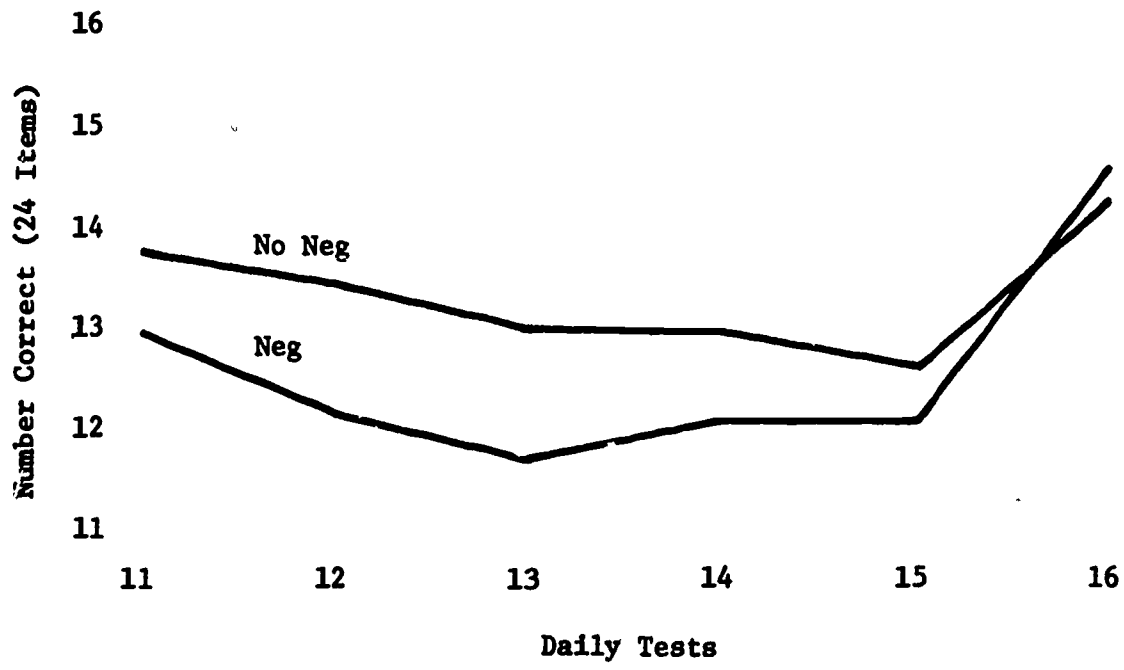

Fig. 10. The average Total Scores for Negative Instance (Neg) and No Neg groups on Daily Tests 11 through 16.

analyzed. The IQ×Fdbk× Neg interaction was also significant on the MC part of Test 13. As before, the Medium and Low groups with Fdbk and Neg did poorly, while the Med and Low with No Fdbk and Neg did well. For No Neg the above was reversed. The High group was superior with Neg and Fdbk, lowest with Neg and No Fdbk. The other interaction, IQ× Sex × Neg, remained close to significance throughout the various ANOVA. In each case the High males did better with Neg than without, and the Low males did poorer with Neg. High females did better with No Neg and Low females poorer with No Neg. IQ × Fdbk was significant (.05) when the Total for Test 11 was examined.

Test 12 On Test 12, IQ was significant at the .05 level on MC and at the .01 level on Comp and Total. IQ × Fdbk was again significant, twice at the .01 level. The means for Total Score were:

Fdbk	High 19.38	Med 9.88	Low 10.25
No Fdbk	High 13.88	Med 13.88	Low 9.88

Sex × Fdbk was significant (.05) for the Comp and Total of Test 12. The means for Total were:

Fdbk	Males 11.17	Females 15.17
No Fdbk	Males 12.92	Females 12.17

An IQ × Sex × Fdbk interaction was significant on the Comp part of 12. The High males did very well with Fdbk and poorly with No Fdbk. The Low also did better with Fdbk, and the Medium males were reversed. All groups of females did slightly better with Fdbk than without.

Test 13 Test 13 had the following significant factors: IQ significant at .01 all three times, Sex (.05 on Comp), Sex × Fdbk (.01 on MC and .05 on Total), Sex × Neg (.05 on MC), IQ × Sex × Fdbk (.01 on MC and .05 on Total), and the IQ × Fdbk × Neg, which was mentioned previously. The means for males and females were 5.25 and 6.42, respectively. The Sex × Fdbk interaction had means as follows:

Fdbk	Males 5.75	Females 7.25
No Fdbk	Males 7.17	Females 6.17

This pattern also held for the ANOVA of the Total. The means for the Sex × Neg interaction were:

Neg	Males 6.67	Females 5.92
No Neg	Males 6.25	Females 7.50

The IQ× Sex× Fdbk interaction may be described as follows: Medium males receiving Fdbk did very poorly. Low males with Fdbk did well. Medium females without Fdbk did poorly while Low females without Fdbk did well. This pattern held for Total at the .05 level.

Test 14 Test 14 showed a significant sex factor (.05) on each ANOVA. For Total Score the means were males 11.21 and females 14.04. The interaction of IQ × Sex × Fdbk was significant. Medium females did better with Fdbk than without, while Medium males did better without than with Fdbk. Note that this is similar to the pattern of the same interaction in Test 13 and Test 12.

Test 15 Fdbk and Sex were significant on Test 15. Males scored 5.88 correct on the Comp, and females scored 7.29. The Fdbk group scored 6.38 on the MC; and the No Fdbk group, 5.33. Neither of these factors was significant on the analysis of the Total Score. IQ × Fdbk was significant for the MC and again for Total. The means for Total were:

Fdbk	High 19.88	Med 11.00	Low 9.00
No Fdbk	High 15.00	Med 12.00	Low 7.62

Fdbk× Neg was another interaction significant at .05 on the MC. The means for MC were:

Fdbk	Neg 5.75	No Neg 7.00
No Fdbk	Neg 5.58	No Neg 5.08

IQ× Sex× Neg was significant at .05 for Total. High males did better with Neg than without, while Low males did much better without Neg than with Neg. Low females did slightly better with Neg than without.

Posttest IIA Pretest II was broken into two parts and each part was analyzed separately. The two parts were called Pretest IIA and IIB. Pretest IIA was exactly the same test as Posttest IIA; and Pretest IIB was the same as Posttest IIB.

Pretest IIA had no significant factors other than the expected IQ difference.

On Posttest IIA Total Score, Sex was significant at the .05 level; males 21.46, females 23.79. On Comp, Fdbk was highly significant (.01); Fdbk, 8.42; No Fdbk, 7.17. The interaction IQ × Fdbk was significant (.01) on MC and again on Total. The means for Total were:

Fdbk	High 32.62	Med 21.75	Low 16.75
No Fdbk	High 23.88	Med 23.62	Low 17.12

Highs did much better with Fdbk than without, while Medium and Low did better without Fdbk.

Test 16 Test 16 was analyzed as part of Experiment 3 since the Lesson 16 Test was unaffected by the Experiment 4 treatment, which was writing exercises given after the test.

Test 16 showed three significant effects: IQ on all three ANOVA at .01, Sex on Comp and Total at .01, and IQ × Fdbk on Comp and Total at .05. On Total Score, males averaged 13.12 and females 15.92. The IQ × Fdbk interaction was as follows:

Fdbk	High 21.50	Med 12.50	Low 10.50
No Fdbk	High 17.50	Med 15.00	Low 10.12

Highs did better with Fdbk, and Medium did better without Fdbk.

Test 17 A Recognition Test was given as part of Test 17. The items to be recognized were transforms learned during Lessons 11 through 16. The Recognition Test showed no significant factors other than IQ as a main effect.

Sums of Tests The Recognition Test score was added to the Test 16 Total and Posttest IIA Total and analyzed. IQ × Fdbk was significant at the .01 level as it had been on Posttest IIA alone. The means were:

Fdbk	High 68.38	Med 38.38	Low 33.12
No Fdbk	High 52.88	Med 47.12	Low 31.88

The Daily Tests were summed together and the following results obtained. IQ was significant at the .01 level and Sex at the .05 level on all three analyses. The Total Score for males was 72.54, and for females, 83.92. IQ × Fdbk was significant (MC, .01; Comp, .05; and Total, .01). The means for Total Score were:

Fdbk	High 117.38	Med 63.88	Low 58.25
No Fdbk	High 93.50	Med 79.12	Low 57.25

The High group again did better with Fdbk, and the Med did better without. One other interaction was significant at the .05 level and this was IQ × Sex × Neg. High and Medium males did better with Neg than without. Low males did much better without. Females were exactly reversed from the males.

When Posttest II scores were added to the sum of the Daily Tests, all the significant effects mentioned for the sum of Daily Tests alone held. IQ, Sex, IQ × Fdbk, and IQ × Sex × Neg kept the same pattern as before.

Pretest IIA compared to Posttest IIA The changes in scores from Pretest IIA to Posttest IIA (these two tests are directly comparable) are shown graphically in Figures 11, 12, 13, and 14. The Low group improved slightly less than the Medium and High. The females improved more than males, and the Fdbk group outdistanced the No Fdbk group. Those who received No Neg showed slightly more improvement than those in the Neg group. The average increase in Total Score for all Ss was 5.9 points (from 16.7 to 22.6 out of a possible 37 points).

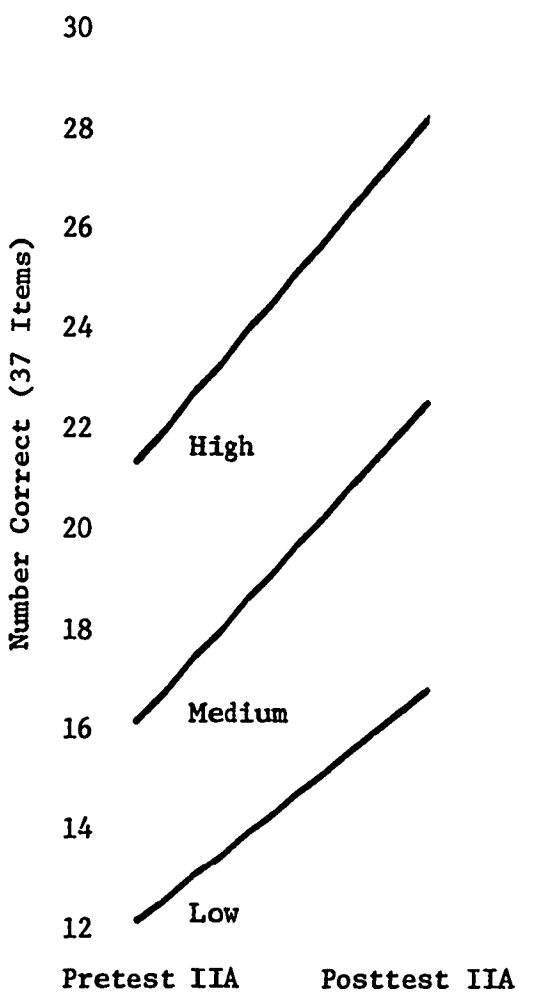

Fig. 11. The average Pretest IIA and Posttest IIA Total Scores for High, Med, and Low IQ groups.

Experiment 4

Forty ANOVA were relevant to the question of whether exercises were helpful in learning transformational grammar. The 40 ANOVA were 3 × 2 × 2 factorial analyses; the factors were IQ, Sex, and Exercises (Ex) or No Ex. Included in the 40 ANOVA were three ANOVA each of Pre-

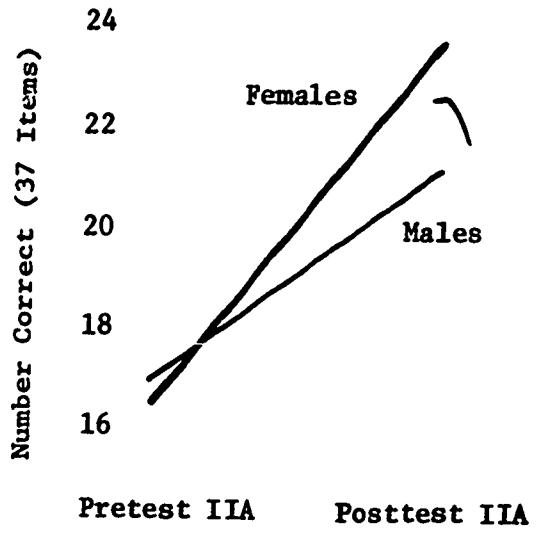

Fig. 12. The average Pretest IIA and Posttest IIA Total Scores for Males and Females.

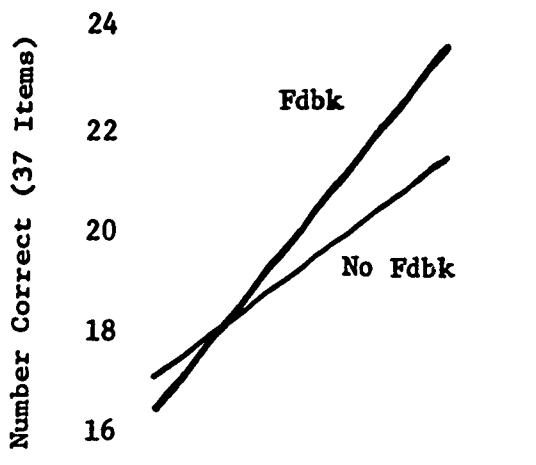

Fig. 13 The average Pretest IIA and Posttest IIA Total Scores for Feedback (Fdbk) and No Fdbk groups.

test IIA and B, Daily Tests 16 through 21, Posttest IIB, Pretest IIB, the sum of Daily Tests 16 through 21, and the sum of Daily Tests plus Posttest IIB. Seven more ANOVA were done on an Exercise Test.

IQ On all ANOVA, IQ (High, Medium, Low) was a significant factor as it had been in all previous analyses. The means and standard deviations for the High, Medium, and Low groups on Total Scores of each test are given in Table 10. Figure 15 is also of interest in noting the pattern of scores for the IQ groups.

Pretest IIB Pretest IIB showed only IQ as a significant factor. Sex and Ex had F ratios less than 1.00. The means for the main effects are given in Table 10. Figure 16 shows the Total Scores of males and females, and Figure 17 shows the Total Scores of the Ex and No Ex groups.

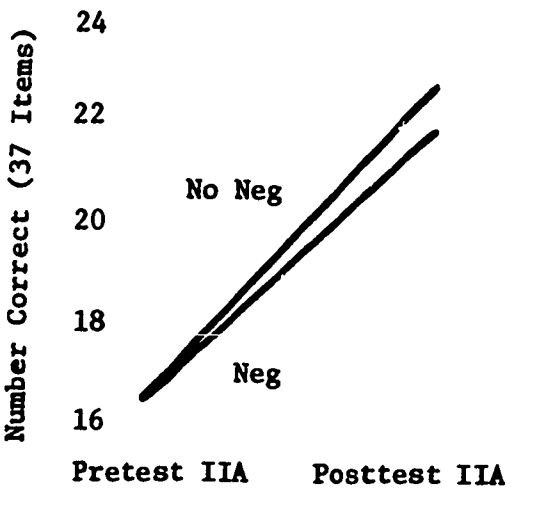

Fig. 14. The average Pretest IIA and Posttest IIA Total Scores for Negative Instances (Neg) and No Neg groups.

Tests 16–21 On Text 16, Sex was significant at the .01 level for Comp and Total Score. Males averaged 13.12 correct answers, and females averaged 15.92 correct out of a possible Total Score of 24.

Tests 17 and 18 showed no significant effects except IQ.

Test 19 results showed a significant sex factor (.05) as Test 16 had shown. The average Total Score for males on Test 19 was 10.21; and, for females, the Total Score was 12.21 correct. On Test 19, Comp part, the interaction of Sex × Ex was significant at .05. The interaction is shown by the means below:

Ex	Males 6.50	Females 6.83
No Ex	Males 5.00	Females 8.08

Males who received Ex performed better than those males who didn't; while females who received Ex did less well than those who received No Ex.

On Test 20, Sex was a significant factor (.05) in the analysis of Total Score. Males scored 10.96 and females 12.67. Ex was not significant nor was any interaction.

Test 21 showed sex significant on all three ANOVA. On Total Score, males averaged 11.08 while females averaged 14.62 ($p < .01$).

Posttest IIB On Posttest IIB, the effects of IQ and Sex were highly significant (.01) on the MC, the Comp, and the Total. On Total, males scored 25.83 and females 31.38 out of a possible 53 points. The interaction Sex × Ex was also significant on all three ANOVA. For Total Score, Sex × Ex reached the .01 level, and the means were:

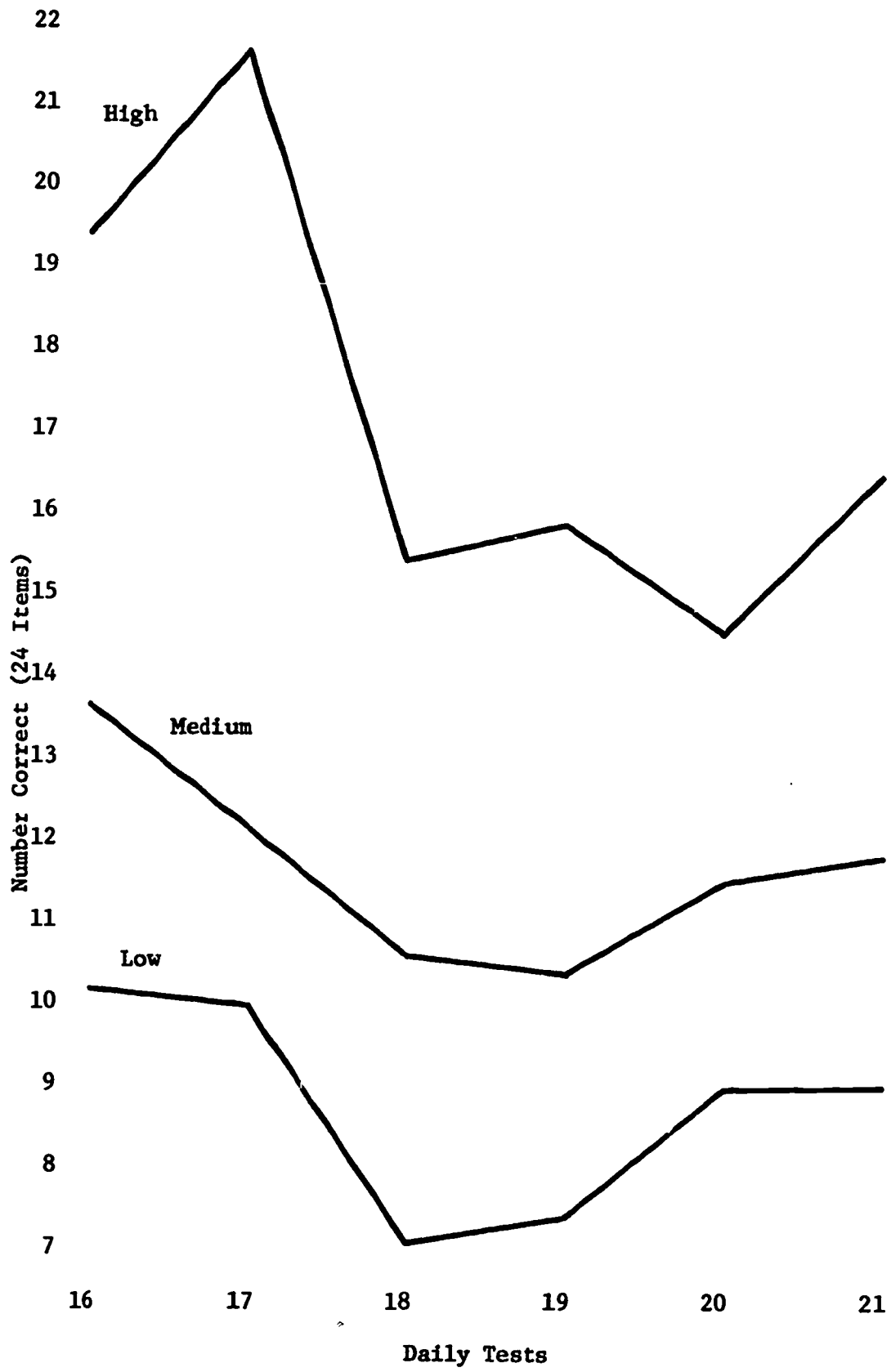

Fig. 15. The average Total Scores for High, Med, and Low IQ groups on Daily Tests 16 through 21.

Ex	Males 28.00	Females 29.33
No Ex	Males 23.67	Females 33.42

Pretest IIB compared to Posttest IIB The improvement from Pretest IIB to Posttest IIB can be read in Table 10 and in Figures 18, 19, and 20. Again the females improved more than the males. There were only slight differences among IQ groups, however. The average increase in Total Score was 7.3 points (from 21.3 to 28.6 out of 53 points).

Sum of tests In the analysis of the sum of the Daily Tests 16-21 (except for Recognition Part of Test 17 which was not included in the sums), IQ and Sex were the only significant factors. When the Posttest III scores were added to the Daily Tests, the pattern held: only IQ and Sex were significant. The means and standard deviations for all the factors are shown in Table 10.

Table 10

Means and Standard Deviations by Groups for Tests 16 through 21, Pretest IIB and Posttest IIB and Exercises

		High	Med	Low	Male	Female	Ex	No Ex
Pretest IIB	Mean	27.25	22.00	14.62	21.00	21.58	21.67	20.92
	St. Dev.	5.6	6.3	4.1	7.3	7.7	6.4	8.5
Posttest IIB	Mean	35.81	28.44	21.56	25.83	31.38	28.67	28.54
	St. Dev.	5.2	6.0	6.6	8.7	7.0	7.5	9.2
Test 16	Mean	19.50	13.75	10.31	13.12	15.92	14.79	14.24
	St. Dev.	3.6	3.2	3.5	5.2	4.7	5.3	5.0
Test 17	Mean	21.75	12.31	10.12	14.88	14.58	15.08	14.38
	St. Dev.	6.7	5.9	4.3	8.3	6.9	8.4	6.7
Test 18	Mean	15.50	10.81	7.31	10.21	12.21	11.50	10.92
	St. Dev.	4.6	4.5	2.1	5.5	4.6	5.6	4.7
Test 19	Mean	16.00	10.44	7.50	10.21	12.42	11.21	11.42
	St. Dev.	2.7	5.1	2.8	5.1	4.9	5.4	4.9
Test 20	Mean	14.75	11.62	9.06	10.96	12.67	12.25	11.38
	St. Dev.	2.5	3.7	2.1	4.0	3.2	3.9	3.5
Test 21	Mean	17.56	11.94	9.06	11.08	14.62	13.04	12.67
	St. Dev.	3.8	4.7	3.9	5.9	4.3	5.9	5.01
Σ (16, 18-21)	Mean	83.31	58.56	43.25	55.58	67.83	62.79	60.62
	St. Dev.	14.4	18.7	11.5	23.6	19.6	24.1	20.9
Σ (Exercises)	Mean	25.38	19.75	13.94	18.54	20.83	22.25	17.12
	St. Dev.	3.8	5.0	7.2	7.1	7.2	5.7	7.6

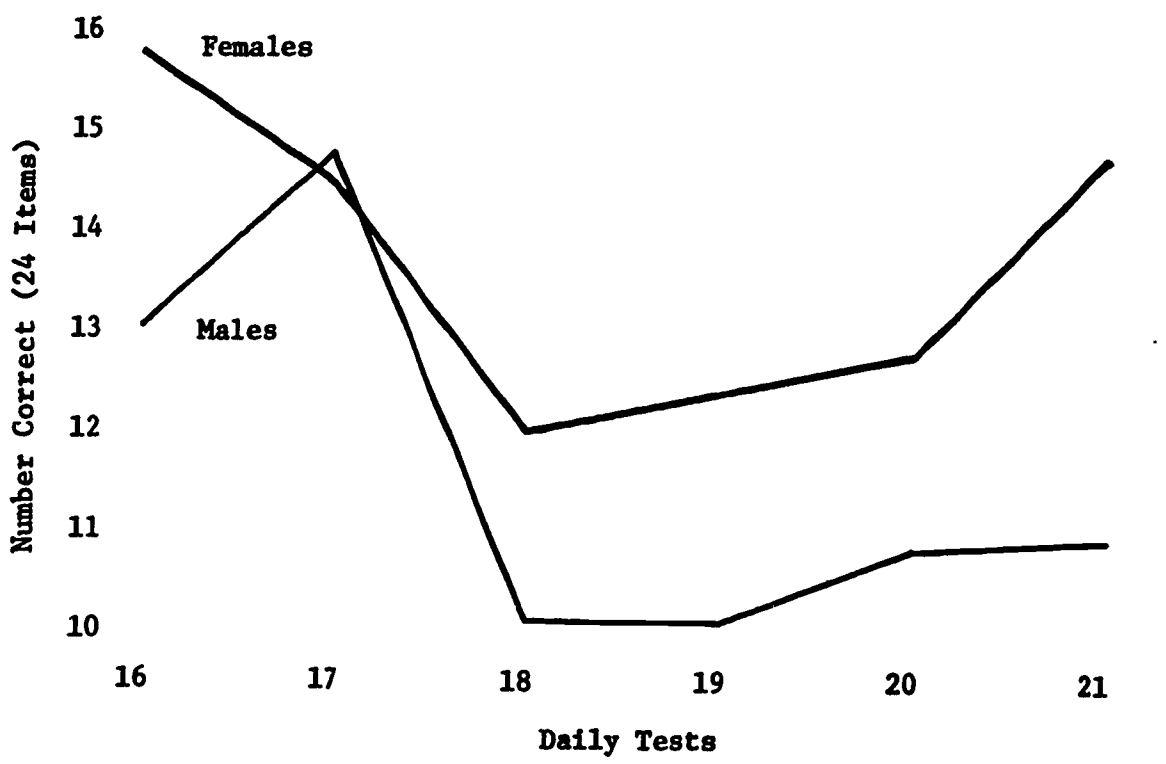

Fig. 16. The average Total Scores for Males and Females on Daily Tests 16 through 21.

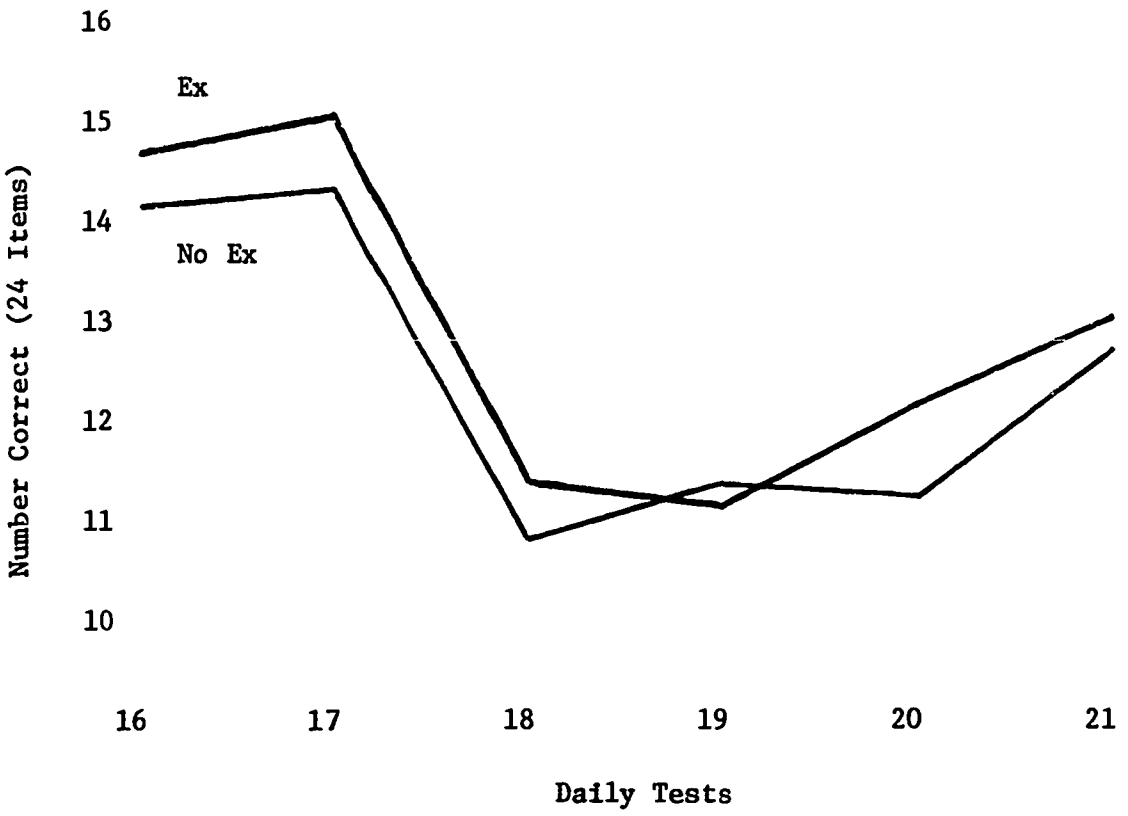

Fig. 17. The average Total Scores for Exercises (Ex) and No Ex groups on Daily Tests 16 through 21.

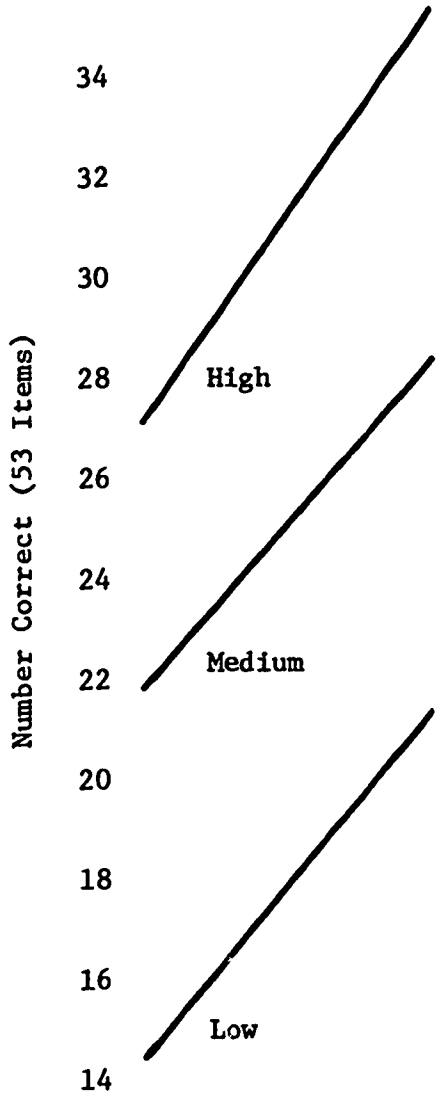

Fig. 18. The average Pretest IIB and Posttest IIB Total Scores for High, Med, and Low IQ groups.

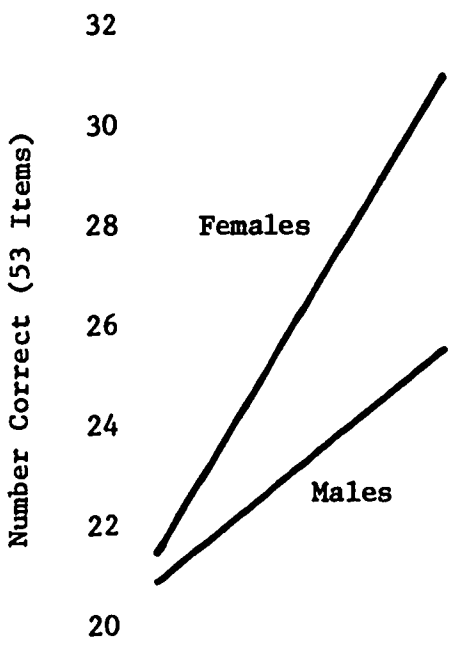

Fig. 19. The average Pretest IIB and Posttest IIB Total Scores for Males and Females.

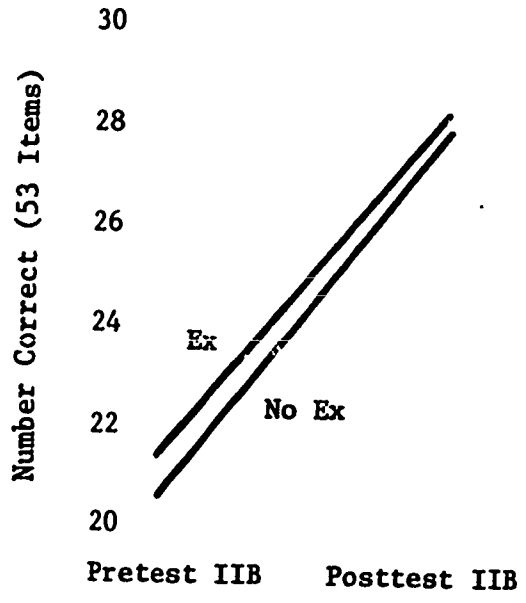

Fig. 20. The average Pretest IIB and Posttest IIB Total Scores for Exercise (Ex) and No Ex groups.

Exercise test The Exercise Test contained four exercises which were much like the exercises half the group had been doing for six lessons. Each of the exercises was analyzed separately by ANOVA. Also the Total Score (sum of the four exercises) was analyzed, together with two other measures, W/T and W/CL. *

Table 11 has the mean, standard deviation, and the significant F ratios for each of these analyses. IQ was significant (.01) on each of Exercises 1-4, and on the Total Score. The means for Total are given in Table 10. Two additional factors were significant for Exercise 1, Sex (.05) and Ex (.01). Males scored 5.17; and females, 6.21. Those receiving Ex as a treatment scored 6.54; those without Ex, 4.83.

On Exercise 2, Sex was not significant. Ex was significant and IQ × Ex was significant, both at the .05 level. The Ex group scored 5.08 and the No Ex group scored 4.17. When this is broken down into IQ groups, the means were:

Ex	High 5.75	Med 5.25	Low 4.25
No Ex	High 5.75	Med 5.25	Low 1.50

For Exercise 3 there was no effect significant except IQ.

Exercise 4 had IQ, Ex, and IQ × Ex significant, all at the .01 level. The mean number of points correct for the exercise group was 5.75 and for No Ex, 4.25. For the IQ × Ex interaction, the means were:

Ex	High 6.62	Med 5.38	Low 5.25
No Ex	High 6.38	Med 5.25	Low 1.12

When the exercises were totaled, IQ, Ex, and IQ × Ex were still significant. The means and standard deviations for this ANOVA are given in Table 10. The Ex group scored 22.25 and the No Ex group 17.12. The means for the IQ × Ex interaction were:

Ex	High 26.25	Med 21.50	Low 19.00
No Ex	High 24.00	Med 18.00	Low 8.88

Two other measures were taken on Exercise 4. These measures were W/T, or words divided by T-unit, and W/CL, or words divided by clauses. In the ANOVA of these two measures, IQ and IQ × Ex were significant in each (.05 level). The means on W/T for High, Medium, and Low were 16.88, 13.44, and 9.62 respectively. The W/T scores for IQ × Ex were:

Ex	High 14.75	Med 11.88	Low 13.62
No Ex	High 19.00	Med 15.00	Low 5.62

For W/CL the means were as follows for IQ:

Ex	High 12.25	Med 7.75	Low 8.75
No Ex	High 11.00	Med 12.62	Low 5.38

For IQ × Ex, the means for W/CL were:

Ex	High 12.25	Med 7.75	Low 8.75
No Ex	High 11.00	Med 12.62	Low 5.38

ITEM ANALYSIS

A total of 441 multiple choice (MC) test items and 338 completion (Comp) items were presented to the Ss over 24 days. Pretests I and II contained 110 MC items and 55 Comp items. All of the pretest items were used again as posttest items, and five of the MC items were used in both Posttest II and Posttest III. Eighteen Daily Tests yielded another 216 new MC items and 228 new Comp items. Each Daily Test contained 12 each of MC and Comp except for Test 17 which had a Recognition Test in place of the MC. An example of daily tests is shown in Appendix E.

Each day's tests (MC and Comp separately) were subjected to the simplified test-analysis procedure described by Stanley (1964). This procedure compares the top 27% of the scores on each test to the bottom 27% to obtain a measure of discrimination (number wrong in the low

*One measure used here was a T-unit (Hunt, 1965). A T-unit is a minimal terminable unit (one main clause plus any clauses attached to or embedded in it). W/T = words/T-unit; W/CL = words/clause.

Table 11

Summary Table of Means, Standard Deviations and F Ratios for Pretest, Daily Tests, and Posttests (Lessons 11–15)

			F Ratios							
	Mean	St. Dev.	IQ	Sex	Ex	IQ × Sex	IQ × Ex	Sex × Ex	IQ × Sex × Ex
Pretest IIA & B MC	22.85	6.76	19.25**						
Comp	12.83	5.60	20.32**						
Total	35.69	11.39	28.20**						
Test 16									
MC	6.71	2.90	20.04**	2.99					
Comp	7.81	2.69	22.85**	10.11**					
Total	14.52	5.09	31.05**	8.42**					
Test 17									
Recognition	8.15	5.60	12.40**						
Comp	6.58	2.66	15.13**						
Total	14.73	7.54	17.44**						
Test 18									
MC	5.29	2.22	15.83**	3.38					
Comp	5.92	3.25	11.72**						
Total	11.21	5.12	16.38**	2.91					
Test 19									
MC	4.71	2.41	16.32**						
Comp	6.60	3.11	17.70**	6.64*			4.30*		
Total	11.31	5.09	22.35**	4.39*			3.45		
Test 20									
MC	5.85	1.86	8.98**						
Comp	5.96	2.42	11.09**	3.58					
Total	11.81	3.66	15.59**	4.21*					
Test 21									
MC	6.40	2.56	6.42**	5.92*					
Comp	6.46	3.49	19.35**	6.53*					
Total	12.85	5.41	18.54**	9.33**					
Posttest IIB									
MC	17.56	4.89	25.52**	12.62**			6.44*	3.07	
Comp	11.04	3.90	24.01**	9.45**			6.37*		
Total	28.60	8.28	32.73**	14.84**			8.56**		
Pretest IIB									
MC	13.67	4.63	15.74**						
Comp	7.62	3.46	15.17**						
Total	21.29	7.43	20.96**						
Σ (MC 16, 18-21)	28.96	9.42	31.21**	7.01*					
Σ (Comp 16–21)	39.33	16.02	22.13**	6.37*					
Σ (Total 16, 18–21)	61.71	22.34	29.65**	8.17**					
Σ (MC 16, 18-21, Posttest IIB)	46.52	13.40	37.41**	11.26**					
Σ (Comp 16–21, Posttest IIB)	50.38	19.49	24.70**	7.57**					

Table 11 (continued)

	Mean	St. Dev.	F Ratios IQ	Sex	Ex	IQ × Sex	IQ × Ex
Σ (Total 16, 18-21, Posttest IIB)	90.31	29.88	33.30**	10.60**			
Exercise 1	5.69	1.93	7.32**	5.63*	15.14**		
Exercise 2	4.62	2.07	16.37**		4.37*		4.37*
Exercise 3	4.38	2.60	10.18**	3.84			
Exercise 4	5.00	2.45	14.68**		8.80**		6.74**
Total Exercises	19.69	7.15	24.59**	2.96	14.81**		3.67*
W/T	13.31	7.80	3.97*				3.46*
W/CL	9.62	4.76	4.66*				3.93*

* $p \leq .05$
**$p < .01$

27% minus number wrong in the high 27% ($W_L - W_H$); and difficulty ($W_L + W_H$) for each item.

For the daily tests, the average item had a discrimination index of 8.59. The average pretest item had an index of 5.75, and on the posttest this average index increased to 7.1. The Comp items discriminated between High and Low scorers more so than the MC items. The average discrimination indexes for Comp items on the daily test, pretests, and posttests were 9.46, 7.60, and 8.82, respectively; and for MC items on daily tests, pretests, and posttests the indices were 7.67, 4.82, and 6.25, respectively.

E. V. Piers (Stanley, 1964) has calculated that for 4-option items responded to by 59 Ss a discrimination index of six or higher is a significant discriminator at the .025 level. Of the 779 items, used in the present program, 539 were significant discriminators at the .025 level. Another 67 items reached a discrimination index of five and thus were nearly significant. These good items included 71 per cent of the MC items and 87 per cent of the Comp items.

QUESTIONNAIRE RESPONSES

In an attempt to get some informal opinions on the programed materials and tests, the students were asked to respond to a questionnaire on the last day of the Summer Project. The questionnaire, which appears in Appendix F, contained 10 questions to be answered "Yes" or "No."

The responses to the 10 questions are listed in Table 12 in terms of per cent of the total population (N = 59). Of particular interest are the responses to Questions 5, 6, and 7. Ss expressed the opinion that the new and unfamiliar concepts did not seem to interfere with what concepts of grammar they already knew. The preference for the programed format over conventional textbook style may imply that the students welcomed a change in instructional style. It also reflects what the students had indicated informally throughout the Summer Project: that they liked working at their own individual pace.

Although some of the vocabulary terms for concepts were already familiar to the students, many of the names for transform processes and resulting structures were necessarily new and, in spite of efforts to simplify wherever possible, were sometimes difficult. In the example sentences, every effort was made to use words which would be within the scope of an eighth grader's vocabulary. It was encouraging to know that the students found both the technical and non-technical words easy to read.

Table 12

Responses to a Questionnaire on the Interest, Readability, Worthwhileness, and Relative Ease or Difficulty of Learning the Programed Materials[a]

Item	% of Ss Answering "Yes"	% of Ss Answering "No"	% of Ss Giving a Response Other Than "Yes" or "No," or Failing to Respond
1. Found Lessons 1-10 (basic sentence patterns) easier than Lessons 11-21 (transform sentences)	83	15	2
2. Revealed response before writing it on answer sheet	79	20	1
a. Revealed response once a lesson - 31% b. Revealed response twice a lesson - 27% c. Revealed response more than twice a lesson - 42%			
3. Would have preferred to have seen the results of all tests	57	42	1
4. Would have liked to have learned more transform grammar	45	52	3
5. Believed that learning sentence patterns and transforms interfered with previously acquired knowledge of English grammar	15	81	4
6. Would have preferred to have the materials written in standard book form	13	86	1
7. Found the vocabulary used in the teaching materials too difficult	1	93	6
8. Found the teaching materials on expanding an NP easier than those replacing an NP	37	62	1
9. Thought it helpful in Lessons 1-10 to have difficult words pronounced	40	57	3
10. Said "noun phrase" when reading the letters "NP"	72	23	5

[a] N = 59.

IV

DISCUSSION AND SUMMARY

PROGRAM MEASURES AND QUESTIONNAIRE RESPONSES

The average time of 21 minutes required to complete the lessons is reasonable and suggests that most lessons can be completed quite comfortably in one class period by students of higher and lower abilities. That the two longest lessons required approximately 30 minutes to complete suggests that some revision and reduction of total length might be necessary. Since Ss needed an average of only 16 minutes to complete the five shortest lessons, the investigators are assured that time need not be a restriction if the insertion of additional frames in subsequent revision seems feasible. In terms of time, lessons ranging from 60 to 75 frames appear most desirable.

An overall error rate of 7.3 per cent does not seem exorbitant, especially if one remembers that most of the content in the lessons was entirely unfamiliar. Although every attempt was made to retain the standard, traditional terms wherever possible, many vocabulary items, as well as the transform processes, were new to the Ss. One would expect that the lessons on transforms, the methods of combining sentences through various processes of reduction and embedding, would necessarily present more difficulties and that these difficulties would show up in the error rate on Lessons 11-21. But, since there is no systematic increase in the overall error rate to correspond with the increase in content difficulty, one may conclude that the Ss were not overwhelmed by the more difficult content. The range of error rates for Lessons 1-10 contains three error rates over eight per cent, the same number as the range for Lessons 11-21. An error rate of 11 per cent for Lessons 3 and 13, apparently the two most difficult lessons, suggests that particular attention need be paid to these lessons in revision. The differences in the error rate for sex or IQ groups were not surprising since the investigators had anticipated that females and the higher IQ group would make fewer errors.

An analysis of the responses to the informal questionnaire administered on the final day of the Summer Project indicated that the Ss found the program enjoyable and that they preferred the programed format over the conventional textbook style.

IQ

In considering the overall results of the experiments, the two stratifying variables included in every experimental design, namely Sex and IQ, showed highly consistent effects. On every test, from the first lesson through the last, the High IQ group scored highest, the Low group scored lowest, and the Medium group scored between the High and the Low groups. In addition to this, the High and Medium groups showed a greater average gain score than the Low group from Pretest I to Posttest I and from Pretest IIA to Posttest II. The High group showed a higher gain than the Medium and Low groups from Pretest IIB to Posttest III. This increased difference between the highs and the other Ss on the later lessons was expected. As was indicated in the description of the content of the lessons, at Lesson 11 a new aspect of the program began. Lessons 1-10 dealt mostly with the basic sentence patterns. At Lesson 11 there was a shift in content from basic sentence patterns to transform rules. The latter half of the program may be considered more complex than the first half. For this reason, it was expected that the Highs would pull away from the other groups because of the increased difficulty of the material. It must be remembered that the IQ range of the high Ss was 126 to 140. Many of these students might be considered intellectually gifted. Clearly, their performance was superior to that of the other Ss on this material. At present we are unable to state what aspects of intelligence lead to superior performance on these materials. Obviously, verbal factors play a large part, but the use of notational devices learned from both

structural and transformational grammar (e.g., NP + V + NP, T-rel) may take abilities more closely associated with non-verbal factors. Further research is required to specify the abilities exemplified in the performance of the High IQ group and to determine whether training procedures will provide students of lower IQ with these abilities.

SEX

The results from the several experiments also indicate that females tended to score higher on the tests than males. The factor tended toward significance on Posttest I and was highly significant on Posttests II and III. On Pretests I and IIA, females as a group scored slightly lower than males. On Pretest IIB, females scored slightly higher than males. On Posttests I, II, and III, however, females scored much higher than males. Thus, in each case, the gain from pretest to posttest was greater for females than it was for males. The general higher performance of females over males on the tests is consistent with research on sex differences in the language arts area.

FEEDBACK

Feedback (Fdbk) tended to be a significant factor in the analyses computed for Experiment 3. In every case where Fdbk was statistically significant, the Fdbk group scored higher than the No Fdbk group. As is indicated in Figure 13, the group receiving the Fdbk treatment scored slightly lower on Pretest II than the group not receiving feedback. By the end of the experiment, the Fdbk group was scoring consistently higher than the No Fdbk group.

It should be emphasized that the term "feedback" as used in this experiment is actually a label for a constellation of treatments administered to the Fdbk group. It includes the feedback of information from the corrected tests of the previous lesson, the positive comments written on the tests, and the discussion held once the tests were in the hands of the students. It was fully realized by the experimenters that the various factors just mentioned were confounded so that information was lost as to which factor had the greatest effect on the Ss' performance. The aim of the feedback treatment was more practical in nature. It was to marshal several learning aids available to the classroom teacher in order to maximize the chances for improving the learning and performance of the students. This aim was achieved with moderate success.

In addition to being significant as a main effect, in several of the analyses the two-way interaction of Fdbk and IQ was significant. In all but one of these interactions, the Med IQ group scored lower than the Low IQ group in the Fdbk condition. This finding was apparently an artifact, however, since scores from Test 11 indicated pre-treatment differences which accounted for the subsequent interactions. The significant interactions were due to an unfortunate (though random) assignment of Ss before the initiation of the Fdbk treatment.

EXERCISES

The performance of the Ss receiving exercises during Experiment 4 has important implications with respect to transfer of training from the program to the student's skill in improving his own writing. On the tests over Lessons 16-21 and on Posttest III, the Ss who received exercises did not score any higher than their controls. It will be recalled that in addition to the final posttest, an exercise test was given. This test consisted of a set of subtests requiring responses very similar to those required by the exercises themselves—responses generated from the application of basic sentence patterns and transform rules. Students did not have the option of selecting some plausible alternative as they did on the MC tests or of supplying a short answer as they did on the completion test. Using the rules, Ss were asked to write several novel sentences.

The analysis of the Total Exercise Test resulted in an F ratio of 14.81 for the exercise factor, with Ss who received exercises on Lessons 16-21 scoring significantly higher ($p < .01$) than those who did not receive them. Comparing the performances on Posttest III and the Exercise Test, it seems reasonable to assume that while the exercises did not help in mastering the content of the lessons, they did aid performance when the information had to be applied in a productive fashion. Put another way, the Ss receiving the exercises not only learned the substantive content but also how to do; the control Ss also learned but were less able to transfer this knowledge to their own writing. All of this implies that in the subsequent revision of the program, exercises should be included wherever the content covered is directly transferable to the student's own writing.

OTHER VARIABLES

Advance organizers, the variable manipulated in Experiment 1, did not reach statistical significance in any of the analyses. The effectiveness of advance organizers should stem from the fact that they are formulated in order to represent a higher level of abstraction and generality than the material which they serve to organize. Two factors may have blocked their effective functioning with the English material. First, the organizers may have been too general and abstract. Concrete examples included in the organizers may have assured both the learning of the organizers and their subsequent effective functioning. A second factor may have been a warm-up effect. The first encounter with material each morning at 8:30 was with the advance organizers. It may have been that the Ss were not able to capitalize on the advance organizers because by the time they had begun to settle down and learn, they were already past the organizers and into the lesson.

The results of Experiment 2 indicated that review did not improve test scores. A highly plausible explanation for this finding is that the review was too brief in relation to the review material already included in the 10 lessons. Lesson 10 consisted of 25 frames which presented new material and 25 frames reviewing material from Lessons 1-10. Thus an average of 2.5 review frames per lesson constituted the review treatment. However, the beginning of each lesson consisted of four or five review frames. The review which all Ss received through these frames may have offset any gains which might have been made through the review treatment.

The use of negative instances in Experiment 3 did not facilitate performance. The only conclusion that it seems possible to draw about adding negative instances to replace positive instances is that the negative instances do not help and probably give slightly less information than the positive instances they replaced. There are several speculations as to the failure of negative instances. The negative instances may not have been as well written as the items they replaced. This could easily have been the case since there were different authors for the two types of instances. Another possible factor is the insertion of the instances. The negative instances were put into the program after the program had been prepared. The negative instances may not have retained the sequence of learning steps. In addition the Neg frames seemed a good deal harder than the positive. Some may have seemed abstruse or irrelevant to the Ss who for 10 lessons had been seeing only positive instances. A last speculation is that negative instances may have helped for some of the concepts but not for others.

Whatever the reason, it does not seem advantageous to insert negative instances of every concept into a program, at the expense of positive instances. When initial learning is the main criterion of effectiveness, certain conditions might produce different results: the negative instances are more carefully written, more appropriately placed, and prepared for in the program sequence; they are used as supplemental material; and they are used only where a concept is clarified and improved by the use of negative instances. Further, it is possible that negative instances might facilitate transfer from the concepts in the present material to other concepts. This too requires further testing.

SUMMARY

Overall, the four experiments utilizing the instructional materials did not give any strong justification for major changes in the materials. Apparently the programed material as it was written worked fairly well without the addition of advance organizers, review, or negative instances. Informative feedback on test results had some good effect and probably should be incorporated in using the programed material in the junior high school. Writing exercises, too, should accompany the instructional materials in order to get students accustomed to the mechanics of writing transforms.

APPENDIX A

ADVANCE ORGANIZER AND CONTROL MATERIAL FOR LESSON I

ADVANCE ORGANIZER

In order to describe something, we usually look at the whole thing and then look at the parts. Knowing the parts and how they fit together helps us in our description. If we are dealing with a number of things, we frequently put them into groups in order to make our description clearer and more organized. During the next two weeks, you will be using this approach in learning to describe English sentences.

One of the first things you will learn is that all sentences may be described in terms of certain basic sentence patterns. <u>There are nine basic sentence patterns</u> in the English language. These nine patterns might be compared to the primary colors that an artist uses. All hues can be obtained from mixtures of red, blue, and yellow which are the three primary colors. Similarly, every sentence you read can be described as taking the form of one of the nine basic sentence patterns, or as a combination or rearrangement of the nine basic sentence patterns. Here are two groups of sentences:

Group 1
<John is president.>
<The boys are a team.>
<All of the men were policemen.>

Group 2
<John is older.>
<The boys are tall.>
<All of the men were happy.>

Although each of these sentences is different, we can classify Group 1 as belonging to one of the nine basic sentence patterns and Group 2 as belonging to another.

Let's look at the sentences again in another way:

<John + is president.>
<The boys + are a team.>
<All of the men + were policemen.>

<John + is older.>
<The boys + are tall.>
<All of the men + were happy.>

These sentences, like every sentence which may be constructed in the English language, can be divided into two main parts. You will learn to call the part to the left of the + sign the <u>subject group</u> and the part to the right of the + sign the <u>predicate group.</u> For this first lesson you will work with the subject group. In all nine basic sentences, the subject group is always a <u>noun phrase.</u> In other sentences, which are rearrangements or combinations of basic sentences, the subject group may or may not be a noun phrase; in basic sentences it is always a noun phrase.

As you learn about noun phrases, you will discover that the last word in all noun phrases is a <u>noun</u>. What is a noun? Rather than depending on the traditional definition of "noun" as the "name of a person, place, or thing," you will learn to use the noun <u>test-sentence</u>. If a word fits in the noun test-sentence, it can be used as a noun. In a later lesson, you will be given other ways which will help you identify nouns.

These pages you have just read are meant to give you a brief overview of today's lesson. Now that you have an idea of what you will be learning, you are ready to begin Lesson 1.

CONTROL MATERIAL

Greek, Roman, and medieval study of Language

During the next five weeks you will be studying your own language. Men have been interested in their languages for a very long

.ime, and many different ideas were developed during this study.

The ancient Greeks in the third century B.C. had one school of thought which believed that the meaning of every word in their language could be traced to the shape of the word. We will illustrate their theory with English examples. The word "blackbird" obviously consists of "black" and "bird." The Greeks reasoned that these birds were named for their color, and indeed, blackbirds are black. But, using the same reasoning the Greeks would have concluded that there was some deep rooted connection between "mushrooms," and "mush" and "rooms." Similarly, they would have looked for a connection between:

lieutenant -- lie, ten, ant
spinach -- spin, itch
capital -- cap, it, all

While this particular approach to language has not survived, other ideas that the Greeks had about their language have lasted for a very long time.

The Romans borrowed many ideas from the Greeks, and it was no different when the Romans came to the study of their own language, Latin. The system of grammar which the Greeks developed was adopted almost in its complete form by the Romans and applied to Latin. Indeed, the Greek generalizations about language were not improved upon until the eighteenth century, when scholars ceased to view language as a direct gift of God, and put forth various theories as to its origin. Two of these later theories were: (1) language began in man's attempts to imitate noises (the "bow-wow" theory); (2) language developed from violent outcries or exclamations (the "ouch" theory).

In the Middle Ages, Latin was changing from its ancient shape into the forms which are commonly known today as the Romance languages—French, Italian, Spanish, and so on. But this change was only occurring in the speech of the various peoples in different geographical areas of Europe. The ancient, classical form of Latin continued to be "The" written language of the educated peoples in these different areas.

Scholars, philosophers, and mathematicians wrote all of their papers in Latin. Many of the Latin terms used today by lawyers are a carry over from the time when all forms of scholarship were conducted in Latin. As the centuries passed, attention was turned from how Latin was to how it ought to be. That is, rules began to be developed as to what was the right way and what was the wrong way to use the language. This approach to the study of language lasted until the 20th century, and is still found in many classrooms in America.

APPENDIX B

ALTERNATE VERSION OF LESSON 15 CONTAINING NEGATIVE INSTANCES IN ONE-FOURTH OF THE FRAMES

Lesson 15: T-ph

1. <Many lizards live under that rock which we just passed.> This is a (basic/transform)sentence.	
2. <The lizards that are living under the rock are harmless.> The transform which produced this sentence is (T-rel/T-pass/T-pos).	transform
3. <My mother gave that man a book.> Is that man a book a related clause?	T-rel
4. <Bats which are flung carelessly may hurt someone.> <Bats flung carelessly may hurt someone.> To make the second transform sentence, we left out the words_____ _____.	No
5. <The diplomat who is on the steps is Danish.> <The diplomat on the steps is Danish.> Here we omitted_____ _____ in the second sentence.	which are
6. Sometimes when adding a related clause we omit some words. The omitted words are a relating pronoun and be. Can words be omitted from the clause in this sentence? <Animals that run in groups are timid.>	who is
7. Remember that is, are, was, and were are all "a form of" be, but to simplify matters we simply call them be. <Othman bought the camel that is drinking at the oasis.> <Othman bought the camel drinking at the oasis.> In this new transform we omit the relating pronoun + __ _____.	No
8. <Othman bought the camel drinking at the oasis.> The underlined group of words is called an -ING PHRASE. <The waitress standing at the counter works days.> Here standing at the counter is an _____ phrase.	be
9. <John, driving quickly, had an accident.> Driving quickly is not a clause, but it is a ph____.	-ing
10. In <You should fix the steps leading to the basement.> the _____ _____ is leading to the basement.	phrase

11. The transform which produces an -ing phrase we will call the PHRASE transform, or T-ph. <Many people riding the elevator were trapped.> This is the result of the phrase transform, which we abbreviate as_____.	-ing phrase
12. Be careful in picking out -ing phrases. Something, king, bingo, sing These words end with -ing but are not considered -ing_____.	T-ph
13. <Logs which are floating downstream will block the river.> <Logs floating downstream will block the river.> T-ph omits the relating prcnoun + _____.	phrases
14. Finish this T-ph transform sentence. <Students who are wearing swimsuits will not be admitted.> <Students_____ swimsuits will not be admitted.>	be
15. To make the transform sentence in frame 14, you omitted the r_____ p_____ + be.	wearing
16. <The miners who were working near the shaft escaped alive.> When we apply T-ph to this transform sentence, we get <The miners_____near the shaft escaped alive.>	relating pronoun
17. <I saw him yesterday when he was there.> When he was there is a related clause. Does it have a relating pronoun to omit?	working
18. <Any hurricanes threatening the coast are watched carefully.> One of the phrases which T-ph produces is the _____phrase.	No
19. threatening the coast We call this an -ing phrase because it is begun by the -ing form of the verb. For the same reason we call a phrase begun by the $past_2$ form of the verb a _____ phrase.	-ing
20. Is this statement correct? All T-ph sentences have an -ing phrase.	$past_2$
21. What is the $past_2$ phrase in the following sentence? <Wallpaper hung upside down may be interesting.>	No
22. Does this sentence contain a $past_2$ phrase? <Wallpaper which is hung upside down may be interesting.>	hung upside down
23. <This village uses water piped in from the Nile.> This transform sentence comes from: <The village uses water_____ _____ piped in from the Nile.>	No
24. Piped in from the Nile is not a ___ phrase.	which (that) is

25. <Dog collars which are buckled too loosely are troublesome.> If we use T-ph with this transform sentence, we get: <Dog collars_____too loosely are troublesome.>	-ing
26. Buckled too loosely is not a(an)_____phrase.	buckled
27. <The skill of the surfers riding the waves amazed us.> <Kites flown too low get tangled in the wires.> Two kinds of phrases which T-ph can produce are _____phrases and_____phrases.	-ing (or prep)
28. T-ph produces a third kind of phrase which expands a NP. You are already familiar with prepositional (prep) phrases. We use them as the completer of Pattern 3 and as the final adverb in other basic patterns. <Some coins in this collection are rare.> What is the prep phrase?	-ing, $past_2$ (any order)
29. <Some coins which are in this collection are rare.> <Some coins in this collection are rare.> Even though T-ph produces three kinds of phrases, the process is always the same. We omit the re. ting pronoun +_____.	in this collection
30. Using T-ph, put the prep phrase immediately after the NP. <All radios that are on this shelf were made overseas.> <All radios_____were made overseas.>	be
31. Are all prep phrases the result of T-ph?	on this shelf
32. Apply T-ph to this transform sentence. <The realtor sold the house that is on the island to Higgins> <The realtor sold the house_____to Higgins.>	No
33. In <The realtor sold the house on the island to Higgins.> T-ph produced a prep phrase. What kind of phrase is not in the following sentence? <The woman on the ladder dusting shelves is our librarian.>	on the island
34. <Flowers picked early may bloom indoors.> T-ph may also produce_____phrases.	$past_2$
35. Sometimes when we have a sentence formed by T-rel, we may wish to make the related clause into a phrase with T-ph. We do this by omitting the_____pronoun + _____.	$past_2$
36. There is a way for us to combine two sentences with T-ph without first doing T-rel. As long as you understand how related clauses are formed (T-rel) and then how phrases are made from the related clauses (T-ph), you will have little difficulty. Try combining these two sentences using T-ph directly. Insert: <A catfish was swimming in the weeds.> Base: <Luke spotted the catfish.> <Luke spotted the catfish_____in the weeds.>	relating, be

37. Swimming in the weeds is an -ing phrase expanding the catfish. Only when a prep phrase expands a NP is it the result of T-ph. Here in the weeds expands swimming, which is not a_____.	swimming
38. Given two sentences, you have a choice as to how you want to combine them. Let's take these two sentences. Insert: <The ferry was moored at the pier.> Base: <The ferry drifted downstream.> If we use T-rel, we get: <The ferry_____was moored at the pier drifted downstream.>	NP
39. Insert: <The ferry was moored at the pier.> Base: <The ferry drifted downstream.> We could also combine these directly with T-ph. If we did that, we would have: <The ferry_____at the pier drifted downstream.>	which (that)
40. Is moored at the pier a prep phrase ?	moored
41. <The ferry moored at the pier drifted downstream.> What NP does the past$_2$ phrase expand ?	No (It contains a prep phrase though.)
42. Use T-ph to combine these sentences directly. Insert: <Seeds are in this packet.> Base: <Many seeds seem dried up.> <Many seeds_____ _____ _____ seem dried up.>	The ferry
43. <Many seeds in this packet seem dried up.> The NP Many seeds is expanded by in this packet, which is a _____phrase.	in this packet
44. Insert: <Seeds are in this packet.> Base: <Many seeds seem dried up.> If we wanted to use T-rel instead of T-ph, our transform sentence would be: <Many seeds_____are in this packet seem dried up.>	prep(prepositional)
45. <That man who was swinging from the beam is a welder.> Suppose you wanted to make the related clause into an -ing phrase. What two words would you leave out ?	which (that)
46. Omitting who was is T-ph. One way to apply T-ph is to omit the relating_____ + _____.	who was
47. We are able to use T-ph in this way only when we already have a transform sentence with a related clause. Could we use T-ph with this sentence ? <The musician is playing the tuba for his neighbors.>	pronoun + be
48. What two words will not be present after we apply T-ph to this sentence ? <The musician who is playing the tuba annoys his neighbors.>	No

49. We can also use T-ph to combine sentences directly without going through T-rel. Insert: <Some words are spoken too crossly.> Base: <Words may damage friendships.> <Words_____too crossly may damage friendships.>	who is
50. Which is not a phrase produced by T-ph ? Noun phrase, prep phrase, -ing phrase, $past_2$ phrase	spoken
51. <That monkey has measles rattling his cage.> Rattling his cage is a misplaced -ing phrase. It should follow the word_____.	noun phrase
52. <The answer guessed by most students appeared sensible.> We have underlined the_____phrase.	monkey
53. <The object in the swamp arrived from Mars.> The third kind of phrase which T-ph can produce is a _____phrase.	$past_2$
54. <The author on stage answering questions wrote a book filled with jokes.> The above sentence has three types of phrases. Do they all expand the same NP?	prep
	No

APPENDIX C
SAMPLE WRITING EXERCISE FOR LESSON 16

Using What You Know #1

1. <The laughing baby crept toward the crystal vase.>

   Underline the two words which were inserted into this sentence with T-BN.

   Write the basic sentences from which the words came.

   a.
   b.

2. Write a sentence of your own which contains a related clause.

   Now write the insert which the related clause was made from.

3. Without using and, but, or or so, use the following sentences for as many different transforms as you can. Next to your sentences write the name of the transform you used, if you know its name.

   <Tony owned a rifle.>
   <Tony cleaned the rifle.>
   <The rifle was .22 caliber.>

4. <Moths flew through the broken window.>

   This sentence was made with T-BN. Write the two sentences which were used for the transform.

5. <A lawn mower was found in the shed.>
   <The lawn mower was rusty.>
   <The shed was rickety.>

   Use any transforms you wish to combine the information in these three sentences. Maybe you can think of several different transform sentences.

APPENDIX D

TOTAL NUMBER OF ERRORS PER LESSON

Ss in Grp	Group	Lesson	1	2	3	4	5	6	7	8	9	10	11
		# Frames	60	50	65	49	63	85	53	68	63	58	61
9	Hi	Male	1	23	37	12	21	34	11	32	26	30	36
11	Hi	Female	3	30	52	9	26	55	22	38	38	41	40
20	Total	Hi	4	53	89	21	47	89	33	70	64	71	76
9	Mid	Male	32	39	58	25	27	48	20	33	27	46	33
10	Mid	Female	13	33	61	13	33	68	24	42	41	47	45
19	Total	Mid	45	72	119	38	60	116	44	75	68	93	78
10	Lo	Male	79	76	130	51	88	107	41	77	71	85	93
10	Lo	Female	58	54	91	42	66	77	40	66	48	61	39
20	Total	Lo	137	130	221	93	154	184	81	143	119	146	132

Lesson	12	13	14	15	16	17	18	19	20	21
# Frames	60	47	41	54	57	62	81	64	46	46
	36	37	29	27	18	27	33	27	19	15
	50	46	23	31	16	25	40	33	17	15
	86	83	52	58	34	52	73	60	36	30
	50	44	26	36	38	25	39	25	16	18
	59	45	46	44	28	20	52	34	19	25
	109	89	72	80	66	45	91	59	35	43
	111	95	59	75	62	64	87	70	42	54
	55	40	49	38	24	31	58	38	21	32
	166	135	108	113	86	95	145	108	63	86

APPENDIX E

SAMPLE TEST ITEMS (LESSON 2)

This is test No. 2. Try to answer as many questions as you can. Guess if you must, but do your best at all times.

Some questions are completion items, and the others are multiple choice. Write your answer in the space provided, or circle the letter of the correct answer as shown below.

0. A dog is:
   a) an animal.
   b) a fish.
   c) a cow.
   d) a verb.

Begin when the instructor tells you.

1. How many types of basic sentences are there ?
   a) 3
   b) 4
   c) 9
   d) 11

2. We put a + sign in a sentence so that:
   a) at least two words are on each side of the +.
   b) the + sign falls in the middle of the sentence.
   c) we separate the subject group from the predicate group.
   d) we have a right and a left side.

3. <The workers were here.> This is a pattern 1, 2, or 3 sentence because it has a form of:
   a) is.
   b) was.
   c) have.
   d) be.

4. <The suit was at the cleaners.> The underlined part is not:
   a) a predicate group.
   b) used as an adverb.
   c) a completer.
   d) a prepositional phrase.

5. Pick out the odd phrase.
   a) was a pilot.
   b) am asleep.
   c) become a man.
   d) is in the house.

6. In basic sentences, which is not true?
   a) The subject group may have only one word.
   b) The subject group is to the right of the verb.
   c) The predicate group may have only two words.
   d) The verb is part of the predicate group.

7. There are_____forms of the word be.
   a) 2
   b) 7
   c) 9
   d) 11

8. In Pattern 1:
   a) the predicate group is NP + be.
   b) the adjective completer follows be.
   c) be is between two NP's.
   d) NP stands for noun phrase completer.

9. Which statement is true?
   a) All Pattern 1 sentences are basic sentences.
   b) All basic sentences are Pattern 1 sentences.
   c) Some Pattern 1 sentences are Pattern 2 sentences.
   d) A Pattern 1 sentence can use the verb ran.

10. Which is a Pattern 2 sentence?
    a) <The guard is tough.>
    b) <Mary is here.>
    c) <Mother is downstairs.>
    d) <John is an engineer.>

11. Structures which follow be are called:
    a) noun phrases.
    b) completers.
    c) adjectives and adverbs.
    d) verbs.

12. The last word in the subject group is:
    a) never followed by a + sign.
    b) a noun.
    c) a noun phrase.
    d) followed by a noun.

13. Write a completer to make this a Pattern 2 sentence.
    <Batman is_____.>

14. Write a Pattern 2 sentence by changing one word in the sentence
    <Five girls are cheerleaders.> _____

15. Jane is a swimmer.
    Two vultures are hungry.
    The structure after be is called a (an)_____

16. Write a Pattern 2 sentence.
    _____

17. The first three kinds of basic sentences have some form of_____as the first word in the predicate group.

18. Write a Pattern 1 sentence.
    _____

19. <It seems very_____.> is the test sentence for a(an)_____.

20. sick
is
the man
a doctor
at the door
Using <u>some</u> of the words from the above list and no others, make a Pattern 1 sentence.

_____

21. Write a sentence of four or five words and underline the predicate group.

_____

22. <The monkey is <u>funny.</u>>
The completer of a Pattern 2 sentence is a(an)_____.

23. The predicate group of Pattern 1 is made up of_____ + _____.

24. Write the test sentence for the noun.

_____

APPENDIX F
QUESTIONNAIRE

_____ _____ _____

Expert Opinion of Student Sample

As we explained at the beginning of this course, you are the experts who will judge the quality of the material you have worked through. Your ideas will help us in revising the material. Please give us your honest answers to the questions below. Do not try to answer them as you think we want them answered—just give us your honest opinion.

1. Were the lessons on basic sentences easier for you than those on transform sentences? Yes No
2. Did you sometimes peek to see what the answer was before writing it on your answer sheet? Yes No
   If yes, how frequently—Once a lesson? Twice a lesson?
3. Would you like to have seen the results of all your tests? Yes No
4. Would you like to learn more about transform grammar? Yes No
5. Does learning sentence patterns and transforms interfere with the grammar you already know? Yes No
6. Would you like the material better if it were written in standard book form? Yes No
7. Were the words used in the program too hard? Yes No
8. Were the lessons on expanding a NP (T-rel, T-ph, T-BN, T-NP) easier than those on replacing a NP (T-NC, T-V_{ing}, T-V_{to})? Yes No
9. Was it helpful to be told the words in each lesson which might give you difficulty? Yes No
10. When you read the letters "NP," did you say to yourself "noun phrase"? Yes No

REFERENCES

Asker, W. Does knowledge of formal grammar function? School and Society, 1923, 17, 109-111.

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune & Stratton, 1963.

Bach, E. An introduction to transformational grammars. New York: Holt, Rinehart, & Winston, 1964.

Blake, R. W. Linguistics and punctuation. English Record, 1964, 15, 9-13.

Boraas, J. Formal English grammar and the practical mastery of English. Unpublished doctoral dissertation, University of Minnesota, 1917.

Bradford, L. P. Study of certain factors affecting English usage. Journal of Educational Research, 1941, 35, 109-118.

Braley, L. S. Strategy selection and negative instances in concept learning. Journal of Educational Psychology, 1963, 54, 154-159.

Buss, A. H. A study of concept formation as a function of reinforcement and stimulus generalization. Journal of Experimental Psychology, 1950, 40, 494-503.

Catherwood, Catherine. A study of relationships between a knowledge of rules and ability to correct grammatical errors and between identification of sentences and knowledge of subject and predicate. Unpublished master's thesis, University of Minnesota, 1932.

Chomsky, N. Syntactic structures. The Hague: Mouton & Co., 1964.

Francis, W. N. The structure of American English. New York: Ronald Press, 1958.

Fries, C. C. American English grammar. New York: Appleton-Century-Croft, 1940.

Glaser, R. (Ed.) Teaching machines and programed learning, II, data and directions, Washington, D. C.: National Education Association, Department of Audiovisual Instruction, 1965, p. 793.

Greeno, J. G. Paired associate learning with massed and distributed repetition of items, Journal of Experimental Psychology, 1964. 67, 286-295.

Hanson, L. F., & Komoski, P. K. School use of programed instruction. In R. Glaser (Ed.), Teaching Machines and Programed Learning, II. Washington, D.C.: National Education Association, Department of Audiovisual Instruction, 1965.

Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-182.

Hoyt, F. S. The place of grammar in the elementary curriculum. Teachers College Record, 1906, 7, 467-498.

Hunt, K. W. Differences in grammatical structures written at three grade levels. Cooperative Research Project No. 1998, United States Office of Education. Tallahassee, Florida: Florida State University, 1964.

Hunt, K. W. Grammatical structures written at three grade levels. National Council of Teachers of English Research Report No. 3. Champaign, Illinois: National Council of Teachers of English, 1965.

Hunt, K. W. Sentence structures used by superior students in grades four and twelve, and by superior adults. Cooperative Research Project No. 5-0313, United States Office of Education. Tallahassee, Florida: Florida State University, 1966.

Huttenlocker, Janellen. Some effects of negative instances on the formation of simple concepts. Psychological Reports, 1962, 11, 35-42.

Johnson, F. S. Structural versus non-structural teaching. College Composition and Communication, 1960, 11, 214-215.

Lees, R. B. The grammar of English nominalizations. Bloomington: Research Center in Anthropology, Folklore, and Linguistics, 1966.

Lin, S. C. Pattern practice in the teaching of standard English to students with a non-standard dialect. Cooperative Research Project No. 1339, United States Office of Education. New York: Bureau of Publica-

tions, Teachers College, Columbia University, 1965.
Mayzner, M. S. Verbal concept attainment: a function of the number of positive and negative instances presented. Journal of Experimental Psychology, 1962, 63, 314-319.
O'Donnell, R. C. The correlation of awareness of structural relationships in English and ability in written composition. Cooperative Research Project No. 1524, United States Office of Education. Mount Olive, North Carolina: Mount Olive College, 1963.
Olson, L. A. Concept attainment of high school sophomores. Journal of Educational Psychology, 1963, 54, 213-216.
Rapeer, L. W. The problem of formal grammar in elementary education. Journal of Educational Psychology, 1913, 4, 125-137.
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.
Roberts, P. Patterns of English. New York: Harcourt, Brace, 1956.
Roberts, P. English syntax. New York: Harcourt, Brace, 1951.
Robinson, Nora. The relation between knowledge of English grammar and ability in English composition. British Journal of Educational Psychology, 1960, 30, Part II, 184-186.
Rogovin, Syrell. Modern English sentence structure. New York: Random House, 1964.
Segal, D., & Barr, Nora. Relation of achievement in formal grammar to achievement in applied grammar. Journal of Educational Research, 1926, 14, 401-402.
Smoke, K. L. Negative instances in concept learning. Journal of Experimental Psychology, 1933, 16, 583-588.
Smoke, K. L. An objective study of concept formation. Psychological Monographs, 1932, 42, No. 4 (Whole No. 191).
Stanley, J. C. Measurement in today's schools. (4th ed.) Englewood Cliffs, N. J.: Prentice-Hall, 1964.
Thomas, O. Transformational grammar and the teacher of English. New York: Holt, Rinehart, & Winston, 1965.
Wallach, M. On psychological similarity. Psychological Review, 1958, 65, 103-116.
Wason, P. C. The processing of positive and negative information. Quarterly Journal of Experimental Psychology, 1959, 11, 92-107.
Whitehall, H. Structural essentials of English. New York: Harcourt, Brace, 1964.
Zidonis, F. J. Generative grammar: a report of research. English Journal, 1965, 54, 405-409.