

© COPYRIGHT

By

Nancy Powell Dixon

1975

THE DEVELOPMENT OF A TEST TO MEASURE THE ABILITY
OF STUDENTS IN TEACHER EDUCATION TO IDENTIFY
COMPREHENSION ACTIVITIES IN TERMS OF
CATEGORIES OF COMPREHENSION SKILLS

By

NANCY POWELL DIXON

Bachelor of Science in Education
The University of Georgia
Athens, Georgia
1965

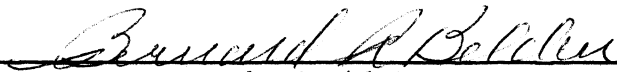
Master of Education
The University of Georgia
Athens, Georgia
1966

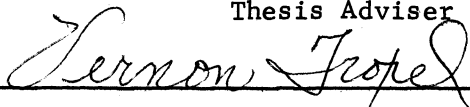
Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF EDUCATION
July, 1975


MAY 12 1976

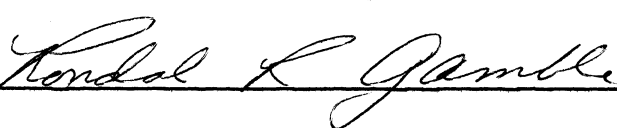
THE DEVELOPMENT OF A TEST TO MEASURE THE ABILITY
OF STUDENTS IN TEACHER EDUCATION TO IDENTIFY
COMPREHENSION ACTIVITIES IN TERMS OF
CATEGORIES OF COMPREHENSION SKILLS


Thesis Approved:



Thesis Adviser








Dean of the Graduate College

938924

ACKNOWLEDGMENTS

Special acknowledgment is made to Dr. Bernard Belden, who suggested and directed this project. I also wish to express appreciation to Dr. Burchard Carr and Dr. Rondal Gamble, members of my advisory committee, and to Dr. Vernon Troxel who not only was a member of the committee but also acted as a judge for the development of this test.

A number of people have been helpful in the development of this instrument. I wish particularly to thank Dr. Barbara Caldwell and Dr. Darrel Ray, who allowed me to use their classes as subjects, and Dr. Charles Smith, Dr. Rita Stuever, and Dr. Marjorie Berends, who acted as judges for the test items.

Appreciation is extended to Dr. Frances A. Scott for her encouragement, example, and inspiration in the pursuit of my graduate work.

Grateful expression is given to my husband, George, for his invaluable assistance and encouragement throughout this study. Finally, I wish to thank my children, Pete and Susan, for their tolerance and for happily filling in the hours between the times I was working on this study.

TABLE OF CONTENTS

Chapter	Page
I. THE PROBLEM.	1
II. REVIEW OF THE LITERATURE	5
III. DESIGN AND CONSTRUCTION OF THE INSTRUMENT.	12
IV. RESULTS AND DISCUSSION	17
V. SUMMARY, CONCLUSIONS, AND IMPLICATIONS	31
A SELECTED BIBLIOGRAPHY	36
APPENDIX A - OUTLINE OF THE BARRETT TAXONOMY.	40
APPENDIX B - A DEVELOPMENTAL TEST ON READING COMPREHENSION SKILLS	42
APPENDIX C - STATISTICAL FORMULAS	51
APPENDIX D - SCORE CHART FOR INSTRUCTED GROUP	55
APPENDIX E - SCORE CHART FOR UNINSTRUCTED GROUP	58
APPENDIX F - DIVISION OF QUESTIONS FOR SPLIT-HALF ANALYSIS.	61
APPENDIX G - ERROR ANALYSIS FOR INSTRUCTED GROUP.	63
APPENDIX H - ERROR ANALYSIS FOR UNINSTRUCTED GROUP.	65

LIST OF TABLES

Table	Page
I. Difference between Mean Scores for Instructed and Uninstructed Groups	18
II. Comparison of Literal Scores of Instructed and Uninstructed Groups	19
III. Comparison of Reorganization Scores of Instructed and Uninstructed Groups	20
IV. Comparison of Inference Scores of Instructed and Uninstructed Groups	20
V. Comparison of Evaluation Scores of Instructed and Uninstructed Groups	21
VI. Comparison of Appreciation Scores of Instructed and Uninstructed Groups	22
VII. Split-Half Reliability.	23
VIII. Undergraduate Grade Point Average for Instructed and Uninstructed Groups	25
IX. Number of Previous Hours in Reading for Instructed and Uninstructed Groups	26
X. Years of Teaching Experience for Instructed and Uninstructed Groups	27
XI. Chronological Age of Instructed and Uninstructed Groups.	27
XII. Comparison between Performance in the Major Areas	29

CHAPTER I

THE PROBLEM

Although reading comprehension is the sine qua non of written communication, little research has been done in this area. Bormuth (1969) in discussing weak areas in reading instruction states, "Comprehension is poor because there is almost no research of any value in the area" (p. 48). This deficiency may be attributed to the fact that the field has been poorly defined, and researchers have not had a common conceptual basis on which to work.

For teachers to be effective in teaching reading, they must have a high level of understanding and competence in the area of comprehension (Smith and Barrett, 1974). In the past, many teachers have had to deal with conflicting points of view concerning the nature of comprehension. Either comprehension was considered a single skill, or it contained so many skills as to be unmanageable. Teacher information regarding comprehension skills is one of the first areas that should be investigated in a systematic attempt to improve instruction in reading comprehension (Wolfe, 1967). In the last decade at least twelve different research projects have investigated teacher knowledge of word recognition skills (Aaron, 1960; Gagon, 1960; Farinella, 1960, Broman, 1962; McCullom, 1964; and Henriksen, 1968). The results of the phonics skills studies revealed that teachers possess a limited knowledge of this area. A similar study has yet to be done in the crucial area of comprehension

skills. If teachers have a limited understanding of phonics skills, it seems likely that they also may be weak in comprehension skills.

The field of reading comprehension has been poorly defined, and therefore, it has been difficult to systematically study it. One of the most recent attempts to define comprehension was completed by Barrett (1968) in his "Taxonomy of the Cognitive and Affective Dimensions of Reading Comprehension". This taxonomy is discussed in Chapter II.

At this point no one has considered the role of the teacher and his knowledge in improving the comprehension abilities of the students. The first step in understanding the teacher's role would be to determine the status of his knowledge in the field. Further, to achieve that, an instrument must be developed.

Statement of the Problem

This study was designed to develop a valid and reliable test to measure the ability of students in teacher education to identify comprehension skills. The test is based on the Barrett Taxonomy of Cognitive and Affective Dimensions of Reading Comprehension.

Hypotheses

The following null hypotheses were formulated for this research.

(1) There will be no difference, significant at the 0.05 level, between the mean scores of the instructed and uninstructed groups on the total test.

(2) There will be no positive correlation, significant at the 0.05 level, between the total test score and the major areas of literal, reorganization, inference, evaluation, and appreciation.

(3) There will be no significant correlation between the split-halves on this test. The 0.05 level of confidence was used in determining whether the hypothesis should be rejected.

Definition of Terms

1. The Barrett Taxonomy, Cognitive and Affective Dimensions of Reading Comprehension - A division of reading comprehension into five skill categories or levels which move from the easy to the difficult. Each level has been further divided into specific kinds of tasks.

2. Comprehension Skills - Those skills defined by the Barrett Taxonomy.

3. Content Validity - "The extent to which a test measures a representative sample of the subject matter and the behavioral changes under consideration" (Gronlund, 1971, p. 78).

4. Construct Validity - "The extent to which a test performance can be interpreted in terms of certain psychological constructs" (Gronlund, 1971, p. 90).

5. Reliability - "This quality implies that a test or instrument consistently measures whatever it measures" (Bledsoe, 1963, p. 85).

Assumptions

There have been several assumptions made in the development of this instrument. These are as follows:

(1) The Barrett Taxonomy can be used as a criterion for discriminating areas of reading comprehension skills.

(2) Students in the reading method courses will be representative of the students in teacher education at Oklahoma State University.

- (3) All of the subjects will be able to read the English language.
- (4) The panel of judges will be competent to discriminate between items in the Barrett Taxonomy.

Limitations

The limitations in this study are as follows:

- (1) The subjects were all enrolled in reading method courses at Oklahoma State University.
- (2) The test does not contain equal numbers of questions in all of the major areas. Nor does it contain questions representing each sub-area of the Barrett Taxonomy. This seeming inconsistency results from this investigator's constraining items to questions found mostly in reading comprehension workbooks.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter presents the literature in three sections. The first section relates to teacher competency in reading. The second part discusses taxonomies of reading comprehension, and the third section presents an investigation based on the Barrett Taxonomy.

Teacher Competency

For many years one of the areas that has been of concern in education is the competency of teachers to give instruction in reading; however, little research has been done in this area. This is particularly true in reading comprehension.

In the 1960's much attention was devoted to teacher knowledge of word recognition skills. Several phonics tests were constructed to obtain data related to this area. The researchers developed the phonics tests to be administered to future teachers or practicing teachers. These studies, which were presented in Chapter I, generally report that teachers have a definite deficiency in their knowledge of phonics principles and generalizations.

Alston (1972) investigated the critical reading ability of classroom teachers in relation to selected background factors such as age, sex, teaching experience, and preparation in reading. A Background Data Sheet was the measure used to obtain the selected background

factors. The findings revealed that chronological age influenced the critical reading skills. As age increased, the test scores of the subjects decreased. Teaching experience was negatively correlated with critical reading test performance. Sex did not appear to be related to test performance, nor did the number of courses taken in the teaching of reading. The differences among the factors and critical reading achievement among the subjects were computed by means of an analysis of variance. A probability of .05 was considered an acceptable level of confidence in variate analysis.

Austin's 1963 report on reading revealed that more than seventy percent of the teachers in the study spent "considerable time" on comprehension; however, after field study observations, it was discovered that quality of teaching did not match quantity. Many teachers failed to teach the basic skills of establishing purposes for reading, relating story content to pupil experiences, providing discussions to clarify misconceptions, identifying pupil strengths and weaknesses, and adapting to individual needs.

Austin's 1961 survey showed that the majority of college courses in reading instruction are geared to the teaching of reading at the primary level and that the middle grade reading skills were mentioned as "marshy areas of instruction." These middle school years are especially important, for this is when the emphasis on higher level comprehension skills would be stressed.

Unless there is training given to teachers during their college years, they can hardly be expected to carry out a successful program of reading with their intermediate grade pupils. Teacher education programs in reading have not been providing teachers with a clear and

consistent conceptualization of reading comprehension, a fact which no doubt has influenced the quality of reading instruction in the classroom.

Even with the wide attention given to Austin's findings, little has been done to present to teachers those comprehension skills which encourage pupils to do more than recognize or recall explicit statements in the children's reading materials. Guszak (1967) conducted an investigation to determine the types of reading comprehension questions that teachers were using. Over a three-day period, he observed and tape recorded all reading groups in four classrooms at the second, fourth, and sixth grade levels. The recordings were then transcribed into written language, and the teachers' questions were analyzed into six classifications according to a Reading Comprehension Question-Response Inventory.

Guszak found several interesting facts. First, the majority of the teachers' questions were of the literal recognition or recall type. Although the percentage of such questions decreased as the grade level increased (78.3%, 64.7%, and 57.8% for grades two, four, and six), it is obvious that the teachers all emphasized literal comprehension during the reading lesson. Secondly, it was discovered that the next most frequently asked questions were of the evaluative type. However, it is important to note that most of these required only a yes or no response. Finally, inferential type questions were third in order of frequency.

Guszak concluded that there was an over-stressing of the literal type questions, especially recall questions concerning minute facts in those classrooms he observed. In addition, he suspected that the inferential and evaluative questions were not used to a larger extent

because the teachers did not have a high enough level of conceptual framework of reading comprehension from which to base their questioning.

Taxonomies

Because the field of reading comprehension is so vast and ill-defined, a number of investigations over the years have attempted to define and limit this area.

Nichols (1948) and Brown (1949) developed a taxonomy for listening comprehension skills. This was based on the distinction between receptive and reflective listening. Receptive listening was defined as those skills primarily associated with accuracy or details in listening, and reflective listening related to skills involving re-organization and inference.

In 1951 Davis developed an outline of nine comprehension skills which provided a basis for the construction of test items to measure the comprehension of high school students. These nine elements of comprehension he listed as follows:

1. Word knowledge.
2. Ability to select the appropriate meaning for a word or phrase in the light of its contextual setting.
3. Ability to follow the organization of a passage and to identify antecedents and references in it.
4. Ability to select the main thought of a passage.
5. Ability to answer questions that are explicitly answered in a passage.
6. Ability to answer questions that are answered in a passage but not in the words in which the question is asked.

7. Ability to draw inferences from a passage about its content.
8. Ability to recognize the literary devices used in a passage and to identify its tone or mood.
9. Ability to determine a writer's purpose, intent, or point of view.

Gray (1960) developed a detailed skills model of reading. He classified common reading activities under four major headings: word perception, comprehension, reaction to what is read, and fusion of new ideas and old. Word perception was defined as being skills for reading. Comprehension was the ability to read the lines, to read between the lines, and to read beyond the lines. Reaction to what is read was identified as critical reading, and fusion of new ideas and old was seen as critical judgement, creative thinking, and combining information with the reader's previous experiences.

Robinson (1966) redefined and expanded Gray's model. She identified word perception as word recognition skills and word meanings. Comprehension was divided into literal and implied meanings. Reaction included intellectual judgements and emotional responses to content. Robinson, however, did not re-define fusion of new ideas and old. She added a fifth aspect of reading, speed and rate. The rate must be flexible and adjusted to the reader's purpose and nature of the material.

Spache (1963) and Cleland (1965) both categorized reading comprehension skills by the mental function involved. Spache's model included five mental processes: cognition (recognition of information), memory (retention of information), divergent production (proceeding from general to specific), convergent production (from part to whole),

and evaluation (critical thinking). Cleland's construct includes perception, apperception, abstraction, appraisal, ideation, and application.

The most extensive and working model of reading comprehension is the Barrett Taxonomy of Cognitive and Affective Dimensions of Reading Comprehension (Clymer, 1968). Barrett based his ideas on the work of Bloom (1956), Sanders (1966), Letton (1958), and Guszak (1967).

The taxonomy is divided into five major skill areas: (a) literal, (b) reorganization, (c) inferential, (d) evaluation, and (e) appreciation. Each of these categories contains examples of specific types of tasks, which are statements of purposes for reading. The five major skills have been arranged in a logical order from the simple level to the more complex level, and the same attempt has been made for the tasks. Moreover, the tasks are cumulative in that performance at one level must utilize all previous levels. The taxonomy in outline form is presented in Appendix A.

In order to improve learning in the area of reading comprehension, a taxonomy such as Barrett developed may have implications for teaching, planning, and evaluating. Smith and Barrett (1974) report that one of the first uses is to assist teachers in determining what reading materials to use or emphasize with respect to comprehension development. Another use is that of designing comprehension tasks. The taxonomy assists in the establishment of purposes for reading a selection before it is read and assists in the development of questions to be asked during or after the reading selection.

An Investigation Based on the Barrett Taxonomy

One of the first research projects using a reading comprehension taxonomy was the work of Albert Cooke (1970). The purpose of his study was to determine the cognitive objectives of the comprehension questions being asked in the manuals, readers, and workbooks of three elementary reading series. The Barrett Taxonomy was used as a basis for analyzing reading comprehension questions. Cooke analyzed 3,536 comprehension questions and found that 55% of the questions were literal, 26% were inferential, 10% were appreciation, 6% were reorganization, and 3% were evaluative type questions. These results tend to reinforce Guszak's study on classroom questioning.

Summary

The few studies concerned with teacher competency have shown that there is a lack of teacher information regarding reading comprehension skills, and this lack may be affecting the quality of reading instruction pupils are receiving. Teachers need to acquire a conceptual framework of these skills in order that effective programs of reading can be developed.

In the evaluation of taxonomies in reading comprehension attention has been given to listening comprehension, the skills involved in comprehension, and the mental functions of comprehension. The Barrett Taxonomy focuses on the practical and is beginning to be used as the basis for research in the area of reading comprehension.

CHAPTER III

DESIGN AND CONSTRUCTION OF THE INSTRUMENT

This chapter presents the design of the instrument and the procedures used in its construction. Information presented includes a description of the instrument and the general procedures followed for development of the instrument.

Description of the Instrument

The test constructed by this investigator was designed to measure the ability of students in teacher education to identify comprehension activities in terms of categories of comprehension skills. The directions stated that the subjects were to select the major area of the Barrett Taxonomy under which each item belonged. These five choices were given at the top of each page in the test. To assist in this task, definitions of each major area were given. The test consisted of 60 items which represented sub-types in each of the five major areas of the Barrett Taxonomy. The number of items per major area was as follows:

- A. Literal - 20
- B. Reorganization - 8
- C. Inference - 15
- D. Evaluation - 9
- E. Appreciation - 8

The major areas of the taxonomy increase in degree of difficulty from literal to reorganization, to inference, to evaluation, to appreciation. Since the Barrett Taxonomy is based on purposes for reading, the areas are not always mutually exclusive, unless the purposes for reading are known. As the purposes were not known in this test, the subjects were instructed to classify a question by the most difficult major area involved.

Development of the Instrument

The Barrett Taxonomy was selected as the reference for the instrument because it defines and limits reading comprehension to a workable area. In addition, it helps in establishing content validity as a representative sample of the subject matter. The taxonomy was discussed in depth in Chapter II.

There were two main sources used in the selection of the test items. The first was reading comprehension workbooks. Twenty-five popular workbooks on levels four, five, and six were selected to be evaluated. All appear on the Oklahoma State Adopted Book List. These five series met the criteria of being current, widely used, and representative. The companies and series titles are as follows:

- (1) Allyn and Bacon, Sheldon Basic Reading Series
- (2) Ginn and Company, The Ginn Basic Readers
- (3) Harper and Row, The Harper and Row Basic Reading Program
- (4) Houghton-Mifflin Company, Reading for Meaning
- (5) The MacMillan Company, The MacMillan Reading Program

Every question throughout the twenty-five workbooks was analyzed. Each of these questions was first categorized and placed under the

appropriate main level of the taxonomy. Next, the question was analyzed to match the sub-type. Barrett's description of the meaning of the main areas and sub-types was the basis for this question analysis and categorization.

Questions for some of the sub-types were not found in the workbook series. Questions in these categories were abstracted from the work of Albert Cooke (1970), who has compiled a nearly exhaustive set of examples of questions relating to the sub-types. A file was made of all the questions analyzed, and items were selected for the test from this composite.

After a large selection of questions had been accumulated, thirty-three items, representing each sub-type, were selected and placed in a random order to make a test. It was then administered to a professional in the field of reading and also to one who knew little about the area. These two persons were instructed to select the major area under which each question belonged, and they were to use no other information to aid their decisions. After the test was scored, the reasoning by which these subjects had arrived at incorrect answers was discussed. This helped to eliminate ambiguous or unclear items. After this elimination, more questions were added and the professional person, again, gave his opinion as to the elimination or rephrasing of certain items. As before, more items were added, and a seventy item test was developed. This test was administered by this researcher to a group of thirteen graduate students in an education class at Oklahoma State University. The purpose of this part of the study was four-fold. The first purpose was to gain feedback from the students concerning the written instructions presented at the beginning of the test. The second purpose was to

eliminate any test items missed by all of the group. The third purpose was to have the class give an opinion as to the method of the test administration. The last purpose was to gain an idea of the time it would take for the test to be administered. Suggestions to improve the written instructions were secured, no items were eliminated by this group, no change was offered in the way of administration, and the students required 40-70 minutes to take the test.

After the pilot test was completed and the group suggestions were applied, the test was then administered to a group of four judges, all of whom had professional knowledge in the area of reading comprehension. The purpose of their screening was to eliminate those test items that did not have a 75% or higher agreement among the judges. Each of the judges was given the test to complete at his leisure, and in addition, each was provided with a skeletal outline of the Barrett Taxonomy along with Barrett's detailed description of the taxonomy. All four judges were instructed to match each question with one of the five major areas. The results indicated that a 75% agreement occurred on sixty of the items, thus ten items were eliminated. The use of this panel of judges was another factor in helping to establish content validity.

Finally, a comparison of two known groups for the purpose of establishing construct validity was made. The sample contained a group of 42 students taking reading method courses at Oklahoma State University during the Fall Semester of 1974. One group of 21 subjects was given five hours of instruction and practice in dealing with the Barrett Taxonomy. In addition to this class instruction, they were given an encompassing outside assignment dealing with the taxonomy. After this exposure to reading comprehension skills, the test was administered by

this researcher in a proctored group test situation with no materials of any kind used to assist the individuals in their responses, and with no discussion of the test taking place prior to or during its administration. The subjects were told that they had as much time as needed to take the test. They used from 40 to 70 minutes. In addition to taking the test, the subjects were asked to provide the following information about themselves: age, degree presently pursuing, degree(s) held, years teaching experience, previous hours in reading method courses, and undergraduate grade point average. The test was then administered in the same manner to a group of 21 students enrolled in another reading method course at Oklahoma State University. This group received no prior instruction on reading comprehension.

The data were collected, all scoring was done by this investigator, and the data were treated statistically. The results are reported in Chapter IV.

Summary

This study was designed to construct a test of reading comprehension to measure the ability of students in teacher education to identify comprehension activities in terms of categories of comprehension skills. The test consisted of 60 multiple-choice items.

The test was administered to one group of 21 students in teacher education who had received instruction in reading comprehension skills, and also, to another group of 21 students who had not received such instruction. All subjects were enrolled in reading method courses at Oklahoma State University during the Fall Semester of 1974.

CHAPTER IV

RESULTS AND DISCUSSION

In this chapter there will be a presentation of the test results, statistical analysis of the data, and a discussion of the data. Other pertinent findings are also included.

Statistical Treatment and Discussion of the Data

The following discussion contains information concerning content validity, construct validity, and internal consistency reliability. The formulas for the statistical procedures are found in Appendix C, and the score charts are summarized in Appendixes D and E.

Content validity was established by the use of the Barrett Taxonomy as a basis for the test construction. Items were selected to represent sub-types in each of the five major areas of the taxonomy. They consisted of questions similar to those found in elementary reading workbooks. Seventy items were evaluated by a panel of four judges, all of whom have professional knowledge in the area of reading comprehension. By using the criterion of 75% or higher agreement with the writer's designation of item category, ten items out of the 70 were eliminated. Thus, a 60 item multiple-choice test was constructed which attained content validity. To establish construct validity, mean scores for the instructed and uninstructed groups were compared using the student's t test, after testing for homogeneity of variances. The level of

significance of the difference was then found using a standard table for a one-tail \underline{t} test. The analysis is presented in Table I.

TABLE I
DIFFERENCE BETWEEN MEAN SCORES FOR INSTRUCTED
AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	60	60
Mean	43.4	35.6
% correct	72.3	59.3
	Standard error	1.90
	F value	2.64
	\underline{t} ratio	4.10
	Degrees of freedom	40

It was found that the mean score of the instructed group exceeded that of the uninstructed group by 4.10 standard errors. Thus, a difference significant at the 0.05 level was achieved.

The test was also examined in terms of the five major areas of the Barrett Taxonomy. Mean scores in each major area for the instructed and uninstructed groups were compared by using the \underline{t} test. The standard error of the difference between the means of the two groups was computed by the same formula as was used to establish validity for the mean scores of the total test. The level of significance of the difference for each major area was then found using a standard table for the one-tailed \underline{t} test.

In order to establish that each major area contributed substantially to the total outcome of the test, a Pearson product moment correlation coefficient was calculated between the scores for the individual areas and the total test. The significance of the correlation was determined by a two-tailed t test.

The results for the literal area are shown in Table II.

TABLE II
COMPARISON OF LITERAL SCORES OF INSTRUCTED
AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	20	20
Mean	14.9	12.6
% correct	74.5	63.0
Standard error	1.02	Correlation coefficient
t ratio	2.25	with total score 0.85
		t ratio 10.2

In the literal area it was found that the mean score of the instructed group exceeded that of the uninstructed group by 2.3 points. This difference is statistically significant. The correlation between the literal area and the total test score was 0.85, which was substantial and was statistically significant.

The results for the comparison of the reorganization scores are found in Table III.

TABLE III
 COMPARISON OF REORGANIZATION SCORES OF
 INSTRUCTED AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	8	8
Mean	6.19	4.67
% correct	77.3	58.3
Standard error	0.37	Correlation coefficient with total score 0.51
<u>t</u> ratio	4.16	<u>t</u> ratio 3.84

In the area of reorganization the mean score of the instructed group was higher than that of the uninstructed group by 1.52 points. This difference was statistically significant. A correlation with the total test score was 0.51, which was statistically significant.

Table IV contains the results of the inference area.

TABLE IV
 COMPARISON OF INFERENCE SCORES OF INSTRUCTED
 AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	15	15
Mean	9.24	7.95
% correct	61.6	53.0
Standard error	0.92	Correlation coefficient with total score 0.80
<u>t</u> ratio	1.39	<u>t</u> ratio 8.64

The major area of inference proved to be the weakest. Here, the mean score of the instructed group only surpassed that of the uninstructed group by 1.29 points, and therefore, was not significant. The correlation between inference and the total test score was 0.80, which was substantial and was statistically significant.

The results for evaluation are presented in Table V.

TABLE V
COMPARISON OF EVALUATION SCORES OF INSTRUCTED
AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	9	9
Mean	6.95	5.71
% correct	77.2	63.4
Standard error	0.47	Correlation coefficient with total score 0.57
<u>t</u> ratio	2.63	<u>t</u> ratio 4.45

The mean score of the instructed group in evaluation exceeded the mean score of the uninstructed group by 1.24 points. This difference was statistically significant. The correlation between evaluation and the total test score was 0.57, which was statistically significant.

Table VI contains the results of the area of appreciation.

The last area of appreciation found the mean scores of the instructed group to be higher than the scores of the uninstructed group by 1.71 points. This difference was statistically significant. A

correlation with the total test score was 0.42, which was, also, statistically significant. This completed the establishment of validity for the test.

TABLE VI
COMPARISON OF APPRECIATION SCORES OF INSTRUCTED
AND UNINSTRUCTED GROUPS

	Instructed	Uninstructed
Number of questions	8	8
Mean	6.14	4.43
% correct	76.7	55.3
Standard error	0.43	Correlation coefficient
<u>t</u> ratio	3.98	with total score 0.42
		<u>t</u> ratio 2.93

In order to determine test reliability, the split-half method was used. The questions on the test were divided by an unbiased coin toss into two equal sections. The sections were then examined to see if they contained equal numbers of questions in the five major areas. As they did not, sufficient questions, again chosen at random, were exchanged between the two halves to achieve this balance. This division is shown in Appendix F.

To relate the scores in the two split-halves, Pearson product-moment correlation coefficients were calculated. These were done for the total group of 42 subjects. The Spearman-Brown formula was used to obtain the reliability of the full test from the reliability of the

split-halves. A correlation coefficient of 0.97 was obtained for the total test. A t score of 25.2 showed that this correlation was significant at the 0.05 level. These results are presented in Table VII.

TABLE VII
SPLIT-HALF RELIABILITY

	First Split-Half		Second Split-Half		r	t
	\bar{x}	s.d.	\bar{x}	s.d.		
Entire Test	18.7	3.30	19.2	2.98	0.97	25.2
Literal	6.81	1.46	6.91	1.69	0.95	19.2
Reorganization	2.43	1.31	3.00	0.63	0.41	2.84
Inference	3.81	2.03	3.50	2.48	0.80	8.43
Evaluation	2.79	2.08	2.98	1.12	-0.004	-0.03
Appreciation	2.91	0.71	2.33	0.90	0.78	7.88

\bar{x} = mean

t = t score

s.d. = standard deviation

r = correlation coefficient

Reliability for each of the major areas was also investigated by the split-half method. The procedure for this was the same as that used for the whole test. A 0.05 level of significance was desired for each area. The correlation coefficient for the literal area of +0.95, inference was +0.80, and appreciation was 0.78, and reorganization was 0.41. Each of these was found to be reliable at the 0.05 level. The

very small negative correlation of -0.004 found for evaluation led this investigator to examine this area further. There are several possibilities for this low correlation. The number of evaluation questions in the test amounted to only nine, as compared with twenty in the literal area. Also, it was the area that produced the second largest number of errors both in the test as a whole and with the instructed group by itself. Therefore, it may be concluded that the test needs more evaluation questions, or that the instructed groups in the future need more training in the evaluation area.

Further, a classification of the evaluation questions by sub-area revealed that the two split-halves originally used did not contain similar populations of sub-area questions, a fact which may be attributed to the small number of total evaluation questions in the test. A new division was constructed to obtain equal sub-area populations in the two split-halves. Through the use of these, a correlation coefficient of 0.77 was obtained, and a t score of 7.63 showed that this was significant at the 0.05 level.

Thus, it appears that in the evaluation area, different skills may be required in different sub-areas in order to classify the question as evaluation successfully. This further emphasizes the need for additional work in the evaluation area.

Other Findings

It was of interest to give attention to what factors in addition to instruction in the Barrett Taxonomy may have been significant in producing high or low scores on the test. No generalizations are drawn from these results. However, they indicate areas which may be important

for further study. The factors which have been considered are:
 (1) demonstrated academic performance, as measured by undergraduate grade point average, (2) previous hours in reading method courses, (3) years of teaching experience, and (4) chronological age. In each instance rank difference correlation coefficients have been computed, and their level of significance examined by a t test. The results of these analyses are presented in Tables VIII, IX, X, and XI.

TABLE VIII
 UNDERGRADUATE GRADE POINT AVERAGE FOR
 INSTRUCTED AND UNINSTRUCTED GROUPS

	R	t	p*
Instructed	+0.55	2.83	yes
Uninstructed	+0.58	3.09	yes
Both	+0.50	3.60	yes

R = Rank difference correlation coefficients

t = t score

p* = Significant at 0.05 level

It was found that there was a strong positive correlation with undergraduate grade point average. There was little correlation between previous hours in reading and test score for the instructed group, but a strong positive correlation between previous hours in reading and test score for the uninstructed group. There was a negative correlation

between years of teaching experience and the score on the test. This led to an examination of whether or not chronological age might be the controlling factor in this case. However, the correlation between test score and chronological age is much less than what might have been expected as a chance occurrence with this number of subjects. The negative correlation between test score and teaching experience may be ascribable to the length of time that these subjects have been away from college.

TABLE IX
NUMBER OF PREVIOUS HOURS IN READING FOR INSTRUCTED
AND UNINSTRUCTED GROUPS

	R	t	p*
Instructed	+0.05	0.20	no
Uninstructed	+0.31	1.43	no

R = Rank difference correlation coefficient

t = \underline{t} score

p* = Significant at 0.05 level

In this same area of additional factors which may have affected scores on the test are degrees held by the subjects. In the instructed group 19 subjects held Bachelor degrees and two held Master degrees. In the uninstructed group 14 had Bachelor degrees, three had Master degrees, and three were undergraduates.

TABLE X
YEARS OF TEACHING EXPERIENCE FOR INSTRUCTED
AND UNINSTRUCTED GROUPS

	R	t	p*
Instructed	-0.43	-2.10	no
Uninstructed	-0.16	-0.70	no

R = Rank difference correlation coefficient

t = \underline{t} score

p* = Significant at 0.05 level

TABLE XI
CHRONOLOGICAL AGE OF INSTRUCTED AND
UNINSTRUCTED GROUPS

	R	t	p*
Instructed	-0.06	-0.25	no
Uninstructed	-0.08	-0.35	no

R = Rank difference correlation coefficient

t = \underline{t} score

p* = Significant at 0.05 level

In order to identify those major areas in which the subjects experienced the greatest difficulty, the numbers of errors made by the subjects were determined in the following way:

(1) A missed question was counted as an error in the area of the question and in the area into which the subject had incorrectly

categorized it. Thus, the total number of errors in Table XII is twice the total number of missed questions.

(2) The raw scores were converted to z-scores to facilitate comparisons between the areas because of the different numbers of questions in each area. The z-scores are presented in Appendixes F and G.

(3) Mean errors in each of the areas were calculated, and the t test was used to check for significance at the 0.05 level. Stated in the null form, the hypothesis was that mean scores in the major areas for each group would not differ at this level of significance.

The results of the t test for the instructed group were as follows:

(1) Literal was greater than inference, but it was less than appreciation.

(2) Reorganization was greater than inference and evaluation.

(3) Inference was less than literal, reorganization, and appreciation.

(4) Evaluation was less than reorganization and appreciation.

(5) Appreciation was greater than literal, inference, and evaluation.

Results of the t test for the uninstructed group showed that appreciation was significantly greater than the other areas. There were no other significant differences. These results are found in Table XII.

TABLE XII
COMPARISON BETWEEN PERFORMANCE IN
THE MAJOR AREAS

Areas	Instructed Group		Uninstructed Group	
	\bar{x}	s.d.	\bar{x}	s.d.
Literal	6.33	3.57	10.05	3.34
Reorganization	4.76	2.67	10.71	3.04
Inference	9.42	4.82	11.90	3.42
Evaluation	7.57	3.68	10.67	3.31
Appreciation	3.80	1.99	4.67	1.70

Significance of Difference between Means

Areas	t	Instructed		t	Uninstructed	
		Sig. at 0.05 level			Sig. at 0.05 level	
L-R	1.57	no		0.34	no	
L-I	2.30	yes		1.57	no	
L-E	1.08	no		0.77	no	
L-A	2.76	yes		6.94	yes	
R-I	3.78	yes		1.29	no	
R-E	2.76	yes		0.42	no	
R-A	1.29	no		7.40	yes	
I-E	1.36	no		0.93	no	
I-A	4.82	yes		8.36	yes	
E-A	4.03	yes		8.60	yes	

\bar{x} = mean
s.d. = standard deviation
t = \underline{t} score
sig. = significant

Total no. of questions in areas

Literal - 20
Reorganization - 8
Inference - 15
Evaluation - 9
Appreciation 8

Summary

In this chapter content and construct validity and internal consistency reliability was established for the test, both as a whole and for the major areas. An examination was made of some of the teacher characteristics which may have affected the outcomes, and also, an identification was made of those major areas in which the subjects experienced the greatest difficulty.

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

This chapter presents a summary of the study. Also included in the chapter are the conclusions drawn from the findings reported in Chapter IV, the educational implications, and the implications for further research.

Summary

This study was designed to construct a valid and reliable test to measure the ability of students in teacher education to identify comprehension activities in terms of categories of comprehension skills.

The null hypotheses tested were as follows:

(1) There will be no difference, significant at the 0.05 level, between the mean scores of the instructed and uninstructed groups on the total test.

(2) There will be no positive correlation, significant at the 0.05 level, between the total test score and the major areas of literal, reorganization, inference, evaluation, and appreciation.

(3) There will be no significant correlation between the split-halves on this test. The 0.05 level of confidence was used in determining whether the hypothesis should be rejected.

The sample used in this study consisted of 42 students enrolled in reading method courses at Oklahoma State University during the Fall Semester of 1974.

The instrument designed by this investigator consisted of 60 multiple-choice items which were in the form of comprehension questions similar to those found in elementary reading workbooks. The subjects were to select the major area of the Barrett Taxonomy under which each item belonged.

Validity for the test was established by determining its content and construct validity. Content validity was established by the use of the Barrett Taxonomy as a basis for constructing the test. Seventy test items were evaluated by a panel of four judges. The requirement that these judges have at least 75% agreement with the writer's designation of item category resulted in the elimination of ten items out of the 70. The final product was a more sophisticated instrument.

Construct validity was determined by comparing mean scores of two known groups. The significance of differences between means was established by the use of the student's t test. The means for the two groups were found to differ significantly at the 0.05 level.

The subjects' performance in each of the individual areas exhibited a positive correlation, significant at the 0.05 level, with the total score on the test. Thus, it was shown that performance in each of the areas contributed substantially to the total outcome of the test.

Reliability for the test was determined by the split-half method. The correlation for the total test was found to be significant. Reliability for each of the major areas was also investigated. Literal, reorganization, inference, and appreciation were found to be significant.

Evaluation was significant when the questions were equally divided by sub-areas. The correlations were tested at the 0.05 level of confidence.

In addition an error analysis was made and a t test was applied. The mean number of errors made by the subjects in each major area was used as a measure of the difficulty experienced by the subjects in correctly identifying the skills connected with each area. The areas ranked in order of increasing difficulty as appreciation, reorganization, literal, evaluation, and inference. The t test showed that at the 0.05 level the difference in the mean scores between adjacent members in this list were not significant, but all differences between non-adjacent members were significant.

Conclusions

On the basis of the data reported, the conclusions which follow have been drawn:

(1) A valid and reliable test was constructed to measure the ability of students in teacher education to identify comprehension activities in terms of categories of comprehension skills.

(2) The test results were educationally significant. Since the instructed group scored higher than the uninstructed group, it is concluded that instruction in the basic comprehension skills is valuable and should, therefore, be taught.

a. Since there was little correlation between previous hours in reading and test score for the instructed group, but a positive correlation between previous hours in reading and test score for the uninstructed group, it would appear

that instruction in reading method courses was helpful to the subjects in being able to successfully perform the tasks required in this test.

Educational Implications

The art of reading is a highly complex process requiring effective functioning of many skills, one area of which is the understanding and application of reading comprehension skills. It seems reasonable to conclude that teachers with limited information regarding these skills are not likely to be able to do an effective job of helping children develop a functional use of these skills. The instrument, it is hoped, will help to alleviate this deficiency.

The test developed in this study has potential for aiding teachers to understand their mastery of reading comprehension skills. The test may be used as a diagnostic tool before instruction or as an evaluative instrument in both pre-service and in in-service teacher education. Special attention needs to be paid to the specific comprehension skills that teachers need in order to be effective.

Implications for Further Research

The need for further research became apparent during the course of this study. Suggestions for additional research in this area of reading are as follows:

- (1) The test may be further refined to include the following additions and changes:

(a) It should contain as an upper limit 100 items to provide a balanced representation of all areas of the Barrett Taxonomy.

(b) It should contain equal numbers of questions in all of the major areas of the Barrett Taxonomy.

(c) It should contain questions for each sub-area of the Barrett Taxonomy.

(2) The test may be used to evaluate the present state of information teachers have of the basic comprehension skills.

(3) The test may be used to determine the effect of other variables such as age, undergraduate grade point average, years of teaching experience, etc., which could be used to draw generalizations about the information of teachers in this area.

(4) A replication of this study at Oklahoma State University would allow for time factor differences.

(5) A replication of this study at other institutions would allow generalizations to be made beyond Oklahoma State University.

(6) The results may be used to compare various instructional schemes, for example, analysis of exercises as found in texts and workbooks against activities involving the formulation of appropriate assessment items.

(7) Status studies may be carried out, for example, pupil achievement and teacher scores on this test.

(8) In order to evaluate the present state of instruction in reading comprehension a content analysis of basic reading courses offered for prospective teachers and experienced teachers at all types of institutions of higher learning may prove to be helpful.

SELECTED BIBLIOGRAPHY

Aaron, Ira E.

- 1960 "What Teachers and Prospective Teachers Should Know About Phonics Generalizations." Journal of Educational Research, LIII (May), 323-330.

Alston, Doris.

- 1972 "An Investigation of the Critical Reading Ability of Classroom Teachers in Relation to Selected Background Factors." Educational Leadership, XXIV (January), 341-343.

Austin, Mary and C. Morrison.

- 1961 The Torch Lighters. Cambridge, Mass.: Harvard University Graduate School of Education.

- 1963 The First R. New York: The MacMillan Company.

Bledsoe, Joseph C.

- 1963 Essentials of Educational Research. Ann Arbor, Michigan: Edwards Brothers, Incorporated, 85.

Bloom, Benjamin S.

- 1956 Taxonomy of Educational Objectives: Handbook I, Cognitive Domain. New York: McKay.

Bormuth, John R.

- 1969 "An Operational Definition of Comprehension Instruction." In K. S. Goodman and J. T. Fleming (eds.), Psycholinguistics and the Teaching of Reading. Newark, Delaware: International Reading Association, Incorporated, 48.

Broman, B. L.

- 1962 "Factors Associated with Teacher Knowledge of Reading Skills." (Unpublished doctoral dissertation, University of Tennessee.)

Brown, J. I.

- 1949 "The Construction of a Diagnostic Test of Listening Comprehension." Journal of Experimental Education, XVIII, 139-146.

Cleland, D. L.

- 1965 "A Construct of Comprehension." In J. A. Figurel (ed.), Reading Inquiry. Newark, Delaware: IRA Conference Proceedings, X, 59-64.

Clymer, T. C.

- 1968 "What Is Reading: Some Current Concepts." Innovation and Change in Reading Instruction. National Society for the Study of Education Yearbook, LXVII, Part II. Chicago, Illinois: University of Chicago Press, 7-29.

Cooke, Albert A.

- 1970 "An Analysis of Reading Comprehension Questions in Basal Reading Series According to the Barrett Taxonomy." (Unpublished Ph.D. dissertation, Cornell University.)

Davis, Frederick B.

- 1951 "Comprehension in Reading." Baltimore Bulletin of Education, XXVIII, No. 3 (January-February), 16-24.

Farinella, John T.

- 1960 "An Appraisal of Teacher Knowledge of Phonetic Analysis and Structural Analysis." (Unpublished doctoral dissertation, The University of Connecticut.)

Gagon, Glenn.

- 1960 "A Diagnostic Study of the Phonetic Abilities of Elementary Teachers in the State of Utah." (Unpublished doctoral dissertation, Colorado State College.)

Gray, William S.

- 1960 "The Major Aspects of Reading." In H. M. Robinson (ed.), Sequential Development of Reading Activities. Supplementary Educational Monographs No. 90. Chicago, Illinois: University of Chicago Press, 13-19.

Gronlund, Norman E.

- 1971 Measurement and Evaluation in Teaching. New York: The MacMillan Company, 78 and 90.

Guilford, J. P.

- 1965 Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company.

Guszak, Frank James.

- 1967 "Teachers' Questions and Levels of Reading Comprehension." In T. C. Barrett (ed.), The Evaluation of Children's Reading Achievement. Newark, Delaware: International Reading Association, 97-110.

Harris, Albert J.

- 1966 The MacMillan Reading Program. New York: The MacMillan Company.

Henriksen, Emmaline B.

- 1968 "An Analysis of Teacher Knowledge of Word Recognition Skills." (Unpublished Ed.D. dissertation, The University of Georgia.)

Hodgman, Charles D.

- 1960 Standard Mathematical Tables. Cleveland, Ohio: Chemical Rubber Publishing Company.

Jones, Dasiy and J. L. Cooper.

- 1964 The Harper and Row Basic Reading Program. New York: Harper and Row.

Letton, Mildred C.

- 1958 "Evaluating the Effectiveness of Teaching Reading." In H. M. Robinson (ed.), Evaluation of Reading. Chicago, Illinois: University of Chicago Press, 76-82.

McCollum, John.

- 1964 "Teachers' Knowledge of Word Analysis Skills and Linguistic Concepts." (Unpublished doctoral dissertation, University of California.)

McKee, Paul and M. Lucile Harrison.

- 1966 Reading for Meaning. Dallas, Texas: Houghton-Mifflin Company.

Nichols, R. G.

- 1948 "Factors on Listening Comprehension." Speech Monograph, XV, 154-163.

Ousley, Odille and David H. Russell.

1966 The Ginn Basic Readers. Boston, Mass.: Ginn and Company.

Ramsey, Z. Wallace.

1962 "Will Tomorrow's Teachers Know and Teach Phonics?" Reading Teacher, XIX (January), 241-245.

Robinson, Helen M.

1966 "The Major Aspects of Reading." In H. A. Robinson (ed.), Reading: Seventy-Five Years of Progress. Chicago, Illinois: University of Chicago Press, 23-32.

Sanders, Norris M.

1966 Classroom Questions. New York: Harper and Row.

Schnepf, Virginia and Odessa Meyer.

1971 Improving Your Reading Program. New York: The MacMillan Company.

Sheldon, William D.

1968 Sheldon Basic Reading Series. Dallas, Texas: Allyn and Bacon, Incorporated.

Smith, Richard J. and Thomas C. Barrett.

1974 Teaching Reading in the Middle Grades. Reading, Mass.: Addison-Wesley Publishing Company.

Spache, George D.

1963 Toward Better Reading. Champaign, Illinois: Garrard Publishing Company.

Weinberg, George H. and John A. Schumaker.

1974 Statistics: An Intuitive Approach. Monterey, Calif.: Brooks/Cole Publishing Company.

Wilkinson, Andrew.

1970 "Research in Listening Comprehension." Educational Research, XII, No. 2 (February), 140-144.

Wolfe, Josephine.

1967 "Applying Research Findings in Comprehension to Classroom Practice." Paper presented at International Reading Association, Seattle, Washington, May, 1967.

APPENDIX A

OUTLINE OF THE BARRETT TAXONOMY

The Barrett Taxonomy: Cognitive and Affective Dimensions of Reading Comprehension*

Literal Comprehension	Reorganization	Inferential Comprehension	Evaluation (Judgement Of)	Appreciation
recognition and recall of:	classifying	inferring of:	reality or fantasy	emotional response to content
details	outlining	details		
main ideas	summarizing	main ideas	adequacy and validity	identification with characters or incidents
sequence	synthesizing	sequence	worth, desirability and acceptability	reactions to author's use of language
comparisons		comparisons	appropriateness	
cause-effect relationships		cause-effect		imagery
character traits		character traits		
		predicting outcomes		
		figurative language		

*Adapted from Barrett's Taxonomy as presented in T. Clymer, "What is Reading?: Some Current Concepts." In H. M. Robinson (ed.), Innovation and Change in Reading Instruction. Sixty-Seventh Yearbook of the National Society for the Study of Education. Chicago: The University of Chicago Press, 1968, 7-29. (Schnepf and Meyer, 1971)

APPENDIX B

A DEVELOPMENTAL TEST ON READING

COMPREHENSION SKILLS

A Developmental Test on Reading
Comprehension Skills

Reading comprehension is a complex of skills that includes reading for literal meaning, for reorganization of ideas, for inference, for making judgements, and for appreciation. The items in this test are comprehension exercises which are commonly found in the workbooks of elementary school readers. In the first half of the exercises student answers are provided, but in the second half only the questions are given.

Directions:

Read the descriptions of these five comprehension skills. Then, select the appropriate skill into which each comprehension exercise falls by marking the correct letter, A, B, C, D, or E. Since the five skills move from the simple to the more complex, one skill is built on the previous one. If there is a question of choice between two skills, select the more difficult one.

A. Literal Comprehension - Exercises for teaching literal comprehension require the reader to recall or locate ideas and information that the author specifically stated in the selection.

B. Reorganization - Exercises for teaching reorganization skills require the reader to analyze, synthesize, and/or organize ideas or information that the author specifically stated in the selection.

C. Inference - In teaching inference skills the reader is required to use his intuition and personal experience to answer questions which demand thinking and imagination that go beyond the printed page.

D. Judgement - In teaching readers to make an evaluative judgement, ideas from the selection are compared with information given by the teacher, other authorities, other written sources, or by the reader's experiences, knowledge, or values.

E. Appreciation - Appreciation is built on the four other comprehension skills. Exercises for teaching appreciation call for the reader to decide the psychological and aesthetic impact the selection has on him.

You will find the five choices as the top of each page. They are as follows:

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

1. Below are some sentences and phrases from the story, "Killer Cat."
Tell what you think the author meant.

a. "a walking nature book"

He knew so much about natural science.

b. "A sure finger on the trigger"

His finger was steady on the trigger.

2. Answer the questions below. Do not reread the paragraphs.

a. What do two long blasts on a train whistle mean?
train about to go What do three short toots mean?
train about to back up

b. In what three ways are messages sent from ship to ship
and from ship to shore? by whistle, radio, and flags

3. The scientists gave specific reasons why some planets are considered
unable to sustain life. Draw a line from the cause in the column
on the left to the effect on the right.

CAUSE	EFFECT
Planet too small.	Too cold.
Planet too far from its star.	Changes in climate
Planet too close to its star.	too extreme.
Planet has too great an orbit.	Oxygen burns off.
	Burned to a cinder.

4. Choose one of the following questions about the Arctic and write a
brief and informative report from information gathered from
encyclopedias and other reference material.

How do people make a living in the arctic region?
What is the Arctic Circle?

5. Write a paragraph telling what might have happened to Roger after
the year 1680. (Pupil's answers will differ.)

6. Check the phrase below that completes the sentence correctly. Do
not turn back to the selection.

In size the Saturn rocket is _____
 taller than the Statue of Liberty
 bigger than a baseball diamond
 400 feet high

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

7. Which story in the unit did you like best of all? Why?
8. After reviewing the main ideas and important details of "Skid Finds His Name in the Newspaper," write a brief summary of the story in your own words.
9. Why was it wrong for Tommy to steal the old man's hen?
10. Grandmother herself was a particular kind of person. Circle the words below that you think describe something about her.

<u>wise</u>	<u>forgiving</u>	selfish
mean	cruel	dull
<u>interesting</u>	<u>warm-hearted</u>	<u>open-minded</u>

11. Read carefully the three reviews of "Castaways in Space" given below. Then, read the three statements labeled a, b, c. Decide on the best description for each review, and mark it a, b, or c.

 c (review) b (review) a (review)
 a. an incorrect review b. A review which includes unimportant details c. A review which includes the most important facts.

12. Number the events in each chapter in the order of the happenings. Do not look back at the selection.

Mystery On Lake Street

<u> 4 </u>	Decoding a letter.
<u> 3 </u>	Bumping into the mailman.
<u> 1 </u>	Finishing all jokes.
<u> 2 </u>	Jim's preparation for fishing.

13. Complete the following outline:

I. Contributions of the Roman builders
 A. The amphitheatre
 B. The dome
 C. The triumphal arch

14. Skim each page listed below and find the word or phrase that belongs in each sentence.

Page 69 Paragraph 1 1. When the Mayflower was in Provincetown Harbor in 1620, whales were seen playing around the ship.

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

15. Use the selection to help you answer each of the following questions:

What besides a top has an axis? the earth

16. Read each statement below and decide whether it is probably a fact or an opinion of the author's. Mark it F or O.

F The men had good equipment.
O They had the best cook in the United States.

17. In each row below underline the topic you would look up in an encyclopedia to find more information about cars.

1. <u>Ford</u>	sailing	trade
2. routes	<u>engine</u>	building
3. <u>automobile</u>	test	wagon

18. Read the following sentences. Decide which are fact and which are fiction.

Fiction 1. The next day the boys decided to look for sunken ships.
Fact 2. Philip IV was king of Spain from 1621-1665.
Fiction 3. Allen found several books written by divers.

19. Write a word or phrase to describe how you would feel if:

You saw a dog hit by a car.
You received just what you wanted for your birthday.

20. Look on page 73. Then follow the directions and answer the questions below.

Was the peak year of French immigration before, during, or after the peak decade for the Germans? before

21. Underline the main idea in each of the following paragraphs.

Dinosaurs were reptiles that lived on earth many millions of years ago, during a period of time called the Age of Reptiles.
This was long before there were any people on earth. No human being has ever seen a living dinosaur.

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

22. Here are four advertisements that might appear in magazines, newspapers, or on television. Read each one carefully and think critically about what it really tells you.

Advertisements given.

- a. Do you think using this product would make you popular if you were not popular before?
b. What is misleading about this ad?
23. Many ideas can be found in "American Whaling" which are not stated by the author but suggested by him. Reread the story to find the hidden ideas which answer the following questions:

Why didn't the cannon balls the Pilgrims fired kill the whales? Their cannon balls weren't powerful enough to shoot the cannon balls through the whales.

Why did the sailors lower the mast and sails when they harpooned a whale? They had to get the ropes and sails out of the way.

24. When writers do not have complete information, or when they are doubtful whether the statements they make can be proved, they often qualify their statements. The qualifying words are underlined in the following sentence:

I have heard that this is the way they get gold in India.

Now, read the four sentences first with the qualifying statements, then omitting the qualifying statements. On the lines below mark each statement true or false as it reads without the qualifying statement. If there is no proof of the accuracy or inaccuracy of the statement, write no proof.

1. false 2. true 3. false 4. no proof

25. You have become well acquainted with the characters in the play, "Caddie and the Indians." Some of them are listed below. Next to each name write the most important thing that you remember about the character.

Kent _____
Hetty _____

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

26. The title, "Three Days to See," is a simple one. Another title could have been one that would have suggested the overall idea that Helen Keller was trying to present. Below are some sayings that may be familiar to you. Circle the one which you think best fits the overall idea of this selection.

Ignorance is bliss.
Beauty is in the eye of the beholder.
Absence makes the heart grow fonder.

27. The job of a page in a kingdom is an important one. Write some of the things that you think a page might have to do.

(Accept any logical response.)

28. Mr. and Mrs. Gay no longer have a lion farm. Where do you suppose motion picture companies and circus owners get or buy their lions now?
29. If you can read between the lines and interpret the facts, you will know how each boy felt. Under each boy's name list the following words which describe the feelings he might have had. You may want to add several.

Tom felt
relieved
curious

Allen felt
brave
thrilled

30. List other parts that are on cars today which may not have been on cars of 1899. fender, taillight, windows, parking lights, etc.
31. Above is a word picture telling how a space pilot may dress when he rockets into space. In the space below, draw the mental picture you see after reading the words. Can you add some interesting details?
32. Find the following information in the story: Date of flight, time in orbit, speed of the space ship, and the height reached.
33. What important thing did Tom find out? You may look back at the story.
34. How do you suppose the explorers conversed with the natives?
35. Read to find out what Jeffery did first.
36. Did the story say if Jon's key and Anne's key were the same?
37. Find the sentence that tells what happened to shorten Mr. Smith's stay at the beach?

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

38. Why did you like or dislike this story?
39. Was it right for Tom to have to pay for the window he broke?
40. Does the author convince you that it is dangerous to hike alone?
41. Would you like to have been in Pete's place when he discovered the missing ship?
42. Read these short selections and then write the best title for each.
43. What do you think happened between the time the landing party crashed and the time they awoke in captivity?
44. Based on the selection, draw a picture that shows how the Indians looked as they crossed the roaring river.
45. List the hardships described in the story that the settlers had to endure.
46. Look in several encyclopedias, and find out how many major earthquakes have been recorded.
47. Did the story say that the two incidents, explosion of the dam and the release of the logs, were related?
48. After reading the paragraph, answer this question. What did the story say that the printing press meant to the world?
49. How does the description of the treatment of the prisoners make you feel?
50. Classify the following people from the story according to their occupations.
51. What happened in the story on the fourth day?
52. After reading the story, list the character traits that you think Micheal Cain possessed.
53. Read the selection, then list the foods that are edible.
54. Do you think Abe Johnson will help them reach the Pacific Coast?
55. In the story, what was the reaction of the king to having his gold disappear?
56. How had the knight shown he was brave?

- (A) Literal (B) Reorganization (C) Inference (D) Judgement
(E) Appreciation

57. Use the main topics and subtopics listed below to outline the article about Death Valley. First fill in the five main topics, then add the subtopics.
58. Put an X before the expression which best explains each of the underlined words or phrases.

John jumped on his skis and took off like a bird.

- began to fly
 X started swiftly away
 flew over the hill

59. Is it really possible for a person to carry two oxen over his shoulders?
60. In what ways do you think Phillip resembles his father?

APPENDIX C

STATISTICAL FORMULAS

The significance of the difference between the means of the two groups were determined by a \underline{t} test. The \underline{t} score was found as follows.

The mean of each group of N subjects was found by

$$\bar{x} = \sum_{i=1}^N x_i / N, \quad (1)$$

where the x's are the scores of the subjects in the group. The standard deviation of each group was found by

$$s = \sqrt{\sum_{i=1}^N (\bar{x} - x)^2 / N}. \quad (2)$$

The standard deviation of the means was found by

$$s_m = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2} \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}, \quad (3)$$

where the subscripts 1 and 2 denote the two groups. The \underline{t} ratio is then given by

$$t = (\bar{x}_1 - \bar{x}_2) / s_m. \quad (4)$$

The level of significance was then determined using standard tables of the one-tailed \underline{t} test.

To determine reliability, Pearson product-moment correlation coefficients were calculated for split-halves of the test. For M questions in each split-half the correlation coefficient is

$$r_{\frac{1}{2}} = \frac{(1/M) \sum_{i=1}^N (x_1^i x_2^i) - \bar{x}_1 \bar{x}_2}{s_1 s_2}, \quad (5)$$

where the subscripts denote the two split-halves and x and s are computed as in equations (1) and (2). The Spearman-Brown formula was used to compute the correlation coefficients for the whole test.

$$r = \frac{2 r_{\frac{1}{2}}}{1+r_{\frac{1}{2}}} . \quad (6)$$

The level of significance of this correlation was determined by computing its t ratio,

$$t = r \frac{\sqrt{2M-2}}{\sqrt{1-r^2}}, \quad (7)$$

and using tables for a two-tailed t test with $2M-2$ degrees of freedom.

To compare the test results with other data on the subjects, a rank-difference correlation coefficient was used.

$$R = 1 - 6 \sum_{i=1}^N (x_i - y_i)^2 / [N(N-1)] . \quad (8)$$

Here, N is the number of subjects and x and y are the ranks of the subjects with respect to the data being correlated. The level of significance of R was found by computing its t ratio

$$t = R \frac{\sqrt{N-1}}{\sqrt{1-R^2}} \quad (9)$$

and using standard tables of the two-tailed t test.

The raw scores of missed questions were converted to Z-scores by the formula

$$z = (x - \bar{x})/s. \quad (10)$$

x = raw score

\bar{x} = mean

s = standard deviation

The F distribution was used to test the null hypothesis that the variances of the distribution of the two samples were the same. F ratios were calculated as follows:

$$F = s_1^2/s_2^2. \quad (11)$$

APPENDIX D

SCORE CHART FOR INSTRUCTED GROUP

Questions
& Correct
Responses

Subjects

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
51 L						R						I									I
52 I								L		A		R	A		L		E		E		
53 R		L	L				L	L								L	L	A			L
54 I						E			E	A		E	E	E	E	E			E	E	
55 L								I				A	E		I	A			I	I	I
56 L		I					I	E		E		I							I	E	I
57 R							L											L			
58 I			E	A	E				L								R	R		A	E
59 E																					I
60 I										E	A	L		A			A		A		E

L = Literal
 R = Reorganization
 I = Inference
 E = Evaluation
 A = Appreciation
 O = Omitted

APPENDIX E

SCORE CHART FOR UNINSTRUCTED GROUP

Questions
& Correct
Responses

Subjects

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 I		L	L	R		L	R		L	R		R		R		R		L			
2 L			I					E	I					I							R
3 L	R	R	R	E	R	R		R	R			R	E		I		R	R			
4 R	E	I	E	L	E		E	E	E	E	I		E	E	E	E					
5 I			R					E			R					A					
6 L		E	E		E													R		E	
7 A															E			E			
8 R					I			I		I		L									L
9 E	A	A	I									L					I				
10 I	E	E	A	E	R	E	E	E		E		E		E			E	E			
11 E			L	I	A	R	R							R							
12 L	R	R	R	R	R	R	R	R	R	R		R			R	R	R		R		
13 R	L	L			L	L	L						L	L		L					
14 L						R	R		R												
15 L			E				I	I		E	E		R	R							
16 E		L	L		I							I		L	I	I	R				
17 R	E	E	I	L	L	I	E	I	E	E	A	I			E	E	I		E		
18 E	R		L	I											I			R			
19 A	I		E		I	I	E		I	I								I	E		
20 L		E	R		R	R						R					R	R	R	R	
21 L		E	R	E	R	I	E	I	R	R	R	R	E			E	I	R	R	R	
22 E	I				R	A						I		A	A		I	A			
23 I			L		L	R	R					E		R	E	R					
24 E	L	R	I		L	I	I	R		I		I	I				R	R			
25 L	E	R	R	E	I	R	A	I	R			E		R	E	A	I	R	R		
26 I	E	E	E		E	E	E		E						E	E	R		E		
27 I			E	E	E	E	E		E	E		R		A		E	E				
28 I			E		E				E			E		E	E	E					
29 A	I		I	I	I	E		I		I	E	E	E		I	I	I	I			
30 I	L		E	E	E	L	L	E				L		R	R		E	E			
31 A	I	L	E	E	I	I		I	I	R	E		I	L		I	I	I	I		
32 L																				R	
33 L										R			R				R		R		
34 I	E		E		E					E	E										
35 L	R	R	R		R	R	R								R						
36 L		E							R	R		I						R			
37 L		R			R	R		R										R			
38 A					E																
39 E		A																A			
40 E	A	A						I	I	A		I		A			A	A	I	A	
41 A	I		E	I	E						I						I				E
42 I	E	R	E	E		E				R	R	R	E	R		E					
43 I			E	R	E		R			E	R									E	
44 A	I	I	L	I	R	I	I		I	I	I	I	I	R	I	R	E	I	I	R	
45 L	R			R	R	R		R				I									R
46 R	L	L	L	E		E	E	E	E	L	L		L					L	L		
47 L		R			R	R	R	R	R	I	I							R	R	E	
48 L												R						R		R	
49 A		E		E	I						I										
50 R		L																I		L	

APPENDIX F

DIVISION OF QUESTIONS FOR SPLIT-HALF ANALYSIS

First Half										
Literal	14	20	21	25	33	37	45	48	51	55
Reorganization		4	8	17	53					
Inference		1	26	28	30	42	43	54		
Evaluation		11	22	39	40					
Appreciation		7	19	38	44					
Second Half										
Literal	2	3	6	12	15	32	35	36	47	56
Reorganization			13	46	50	57				
Inference			5	10	23	27	52	58	60	
Evaluation			9	18	24	59				
Appreciation			29	31	41	49				

APPENDIX G

ERROR ANALYSIS FOR INSTRUCTED GROUP

Z-Scores*

Subjects	Areas				
	Literal	Reorganization	Inference	Evaluation	Appreciation
1	-1.21	-1.78	-1.75	-1.51	-1.41
2	-0.93	-0.66	-1.75	-2.06	-1.41
3	-1.49	-1.03	-1.54	-1.24	+0.10
4	-0.65	-0.28	-0.92	-0.70	+0.10
5	-1.49	+0.46	-0.50	-0.15	+0.60
6	-0.65	-0.28	+0.12	-0.70	-0.40
7	+0.46	+1.59	-0.92	-1.24	-1.41
8	+0.18	-0.66	-0.09	+0.12	-0.40
9	-0.93	-0.28	-0.09	+0.12	+1.11
10	-0.37	-0.66	+0.54	+0.66	-0.90
11	-0.37	-1.03	+0.54	+0.66	+0.60
12	+1.31	-1.41	+0.74	-0.70	+0.10
13	+0.47	+0.46	+0.95	-0.15	-1.41
14	-0.65	-0.66	+0.74	+0.66	+0.60
15	-0.37	+0.46	+0.12	+1.48	+0.10
16	+1.31	+0.09	+0.12	-0.15	+0.60
17	+0.47	+1.59	-0.50	+1.20	+1.11
18	-0.37	+1.21	+0.54	+1.20	-0.40
19	+2.15	+1.59	+1.59	+0.66	+1.61
20	+1.03	+1.21	+2.20	+1.75	+2.11
21	+0.47	+0.09	-0.09	+0.12	-0.90

*A positive z-score indicates a high number of errors.

APPENDIX H

ERROR ANALYSIS FOR UNINSTRUCTED GROUP

Z-Scores*

Subjects	Areas				
	Literal	Reorganization	Inference	Evaluation	Appreciation
1	-0.91	-0.56	+0.61	+1.01	+1.37
2	-0.91	-0.89	-1.14	-0.81	-1.57
3	-0.01	-0.23	-1.73	-1.11	+0.10
4	+2.08	-0.95	-0.26	+1.91	+0.78
5	+1.18	-0.89	+0.61	+0.40	-1.57
6	-0.91	-0.89	+0.91	+1.01	-0.39
7	-1.21	-1.55	-0.26	+0.40	+1.37
8	+1.78	-0.75	+1.49	+0.10	+0.10
9	-0.31	+0.42	+0.32	+0.70	-0.98
10	+0.28	-0.75	+0.03	-0.20	-0.39
11	+1.48	+3.06	+0.61	-0.50	-0.39
12	-1.81	-0.95	-0.85	-2.62	+0.78
13	-0.31	-0.56	+0.03	-0.81	-0.39
14	-0.58	-1.22	+1.20	+0.40	-0.98
15	-0.31	-0.75	+0.85	+0.40	+0.10
16	-1.21	-0.23	-1.43	-1.71	-0.98
17	+0.28	-0.23	+1.20	+0.70	-0.39
18	-0.01	-0.23	+1.20	+1.01	+2.55
19	-0.01	+1.08	+0.61	-0.20	+0.78
20	-0.58	+1.08	-0.26	+0.10	-0.39
21	-0.31	-0.56	-2.02	-0.20	+0.10

*A positive z-score indicates a high number of errors.

VITA

Nancy Powell Dixon

Candidate for the Degree of

Doctor of Education

Thesis: THE DEVELOPMENT OF A TEST TO MEASURE THE ABILITY OF STUDENTS
IN TEACHER EDUCATION TO IDENTIFY COMPREHENSION ACTIVITIES IN
TERMS OF CATEGORIES OF COMPREHENSION SKILLS

Major Field: Elementary Education

Biographical:

Personal Data: Born in Villa Rica, Georgia on October 22, 1943,
the daughter of Mr. and Mrs. Jack W. Powell; married George
S. Dixon, Jr. on June 8, 1966; two children; Peter, 1969 and
Susan, 1972.

Education: Graduated from Buford High School, Buford, Georgia in
June, 1961; received Bachelor of Science in Education degree
from the University of Georgia in June, 1965; received the
Master of Education degree from the University of Georgia in
August, 1966; completed requirements for the Doctor of Educa-
tion degree at Oklahoma State University in July, 1975.

Professional Experience: Special education teacher, Gracewood
State School and Hospital, Gracewood, Georgia, Summer, 1965;
Clinical assistant (supervised practicing teachers in special
education) at the University of Georgia, Summer, 1966; Public
school special education teacher, Athens, Georgia and Cooke-
ville, Tennessee, 1966-1968; Remedial reading teacher, Daniel
Arthur Rehabilitation Center, Oak Ridge, Tennessee, 1968-1970;
Remedial reading teacher at elementary school in Morrison,
Oklahoma, 1974-present.