1978

A comparative study of the effects of two approaches to teaching reading comprehension on achievement of fourth and fifth grade students of middle and low socioeconomic status

Alfred John Gordon

University of the Pacific

Follow this and additional works at: https://scholarlycommons.pacific.edu/uop_etds

Part of the Education Commons

Recommended Citation


https://scholarlycommons.pacific.edu/uop_etds/3192

This Dissertation is brought to you for free and open access by the Graduate School at Scholarly Commons. It has been accepted for inclusion in University of the Pacific Theses and Dissertations by an authorized administrator of Scholarly Commons. For more information, please contact mgibney@pacific.edu.
A COMPARATIVE STUDY OF THE EFFECTS OF TWO APPROACHES TO TEACHING READING COMPREHENSION ON ACHIEVEMENT OF FOURTH AND FIFTH GRADE STUDENTS OF MIDDLE AND LOW SOCIOECONOMIC STATUS

A Dissertation
Presented to
the Graduate Faculty of the
University of the Pacific

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

by
Alfred John Gordon
February 1978
This dissertation, written and submitted by

ALFRED JOHN GORDON

is approved for recommendation to the Committee on Graduate Studies, University of the Pacific:

Oscar J. Jarvis
Dean, School of Education

Dissertation Committee:

Michael B. Gilbert
Chairman

B. R. Hopkins

Juanita E. Curtis

Randale Romberg

A. L. Hutchins

Dated February 7, 1978
A COMPARATIVE STUDY OF THE EFFECTS OF TWO APPROACHES TO TEACHING READING COMPREHENSION ON ACHIEVEMENT OF FOURTH AND FIFTH GRADE STUDENTS OF MIDDLE AND LOW SOCIOECONOMIC STATUS

Alfred John Gordon

ABSTRACT

The Problem

The problem was divided into two parts. The first part investigated the question whether students who were taught reading comprehension skills by the PIRAMID objective-based approach showed a greater gain than students who were taught the same skills by use of the basal reader approach. The second part was to ascertain the degree of concordance between results of the criterion-referenced test (CRT) and results of the norm-referenced test (NRT).

The Purpose

The purpose was twofold: (1) to compare gains in reading comprehension achievement of an experimental group with gains of a control group, and (2) to compare the pretest and posttest results obtained from the CRT with those obtained from the NRT to determine the extent of relationship between the two measures.

Methodology

The study was conducted in thirty-four fourth and fifth grade classrooms in three school districts in Northern California with a sample of 670 students participating. An experimental group consisted
of 369 students and a control group consisted of 301 students. Of the total sample of 670, 407 were classified as students of low socioeconomic status (SES) and 263 were classified as students of middle SES. The experimental group teachers taught reading comprehension by the PIRAMID objectives-based approach and the control group teachers taught reading comprehension by the basal reader approach. The two groups were assigned to a Nonrandomized Pretest and Posttest Control Group Design. Pretests and posttests on the CRT and the Stanford Achievement Test were administered in one school district and pretests and posttests on the CRT and the Comprehensive Tests of Basic Skills were administered in the other two school districts. Analyses of covariance procedures were used to test Hypotheses 1-6 and a Pearson Correlation analysis was used to test Hypotheses 7 and 8.

Findings

(1) The experimental group showed a significantly greater gain than the control group as evidenced by both measures. (2) Fourth grade subjects showed a significantly greater gain than fifth grade subjects as evidenced by the CRT. (3) Subjects of middle SES showed a significantly greater gain than subjects of low SES as demonstrated by both measures. (4) The experimental approach was more effective for fourth grade and middle SES subjects as evidenced by an approach by grade and an approach by SES interaction on the CRT. No significant differences occurred when grade or interactions were taken into consideration on the NRT. The correlation between results of the CRT and the NRT was substantially high and consistent across all measurements.
Conclusions

It was concluded that the PIRAMID objectives-based approach was more effective than the basal reader approach in teaching reading comprehension skills to fourth and fifth grade students of low and middle socioeconomic levels. The CRT was highly comparable to the NRT in assessment of reading achievement.

Administrative Implications

This study has implications for the elementary school principal relating to staff development programs, alternative school organizational patterns, the use of instructional objectives to provide balance in the total curriculum, the establishment of a resource center, parent education, budgetary provisions, and evaluation.

Recommendations

The following recommendations are presented: (1) to replicate this study with a larger sample size and a broader range of SES groups; (2) to conduct a study similar to the current study in which the amount of investigator supervision would be reduced; (3) to conduct a similar study on the PIRAMID math instructional system in the subtest area of math concepts; (4) to follow up the current sample in this study to determine how lasting would be the effects of the experimental approach; (5) the PIRAMID Consortium should seek ways to reduce the amount of testing involved with its Instructional System; (6) establish grade equivalent norms on the PIRAMID CRT's; and (7) the California State Department of Education should refine its method of determining SES by including other indicators rather than relying solely on parents' occupation.
ACKNOWLEDGMENTS

The investigator wishes to express his thanks to the many individuals who made the completion of this study possible. I am particularly grateful to my Dissertation Committee, Dr. Michael B. Gilbert, Chairman, Dr. Bob R. Hopkins, Dr. Juanita G. Curtis, Dr. Randall E. Rockey, and Dr. Alan L. Mikels, for their counsel, guidance, and critical analyses during the preparation and development of the study.

I gratefully acknowledge the cooperation and support of Dr. Wayne Sorenson, Director of Program Development and Research for the Hayward Unified School District; Dr. Bob Rothchild, Superintendent of the Pittsburg Unified School District; and Mr. Andros Karperos, Superintendent of the Yuba City Unified School District.

Appreciation is expressed to participating students, teachers, and principals of the three school districts, and especially to Mrs. Paula McIntire, Chairperson of the PIRAMID Consortium at Yuba City. Without their cooperation and involvement, the study would not have been possible.

Finally, my family deserves the greatest thanks—my wife, Mildred, my daughter, Kelly Gail, and my sons, Gregory Jay and Marc Spencer—for their continuing assistance and encouragement.

A.J.G.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES AND CHARTS</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. BACKGROUND OF THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>The Problem</td>
<td>5</td>
</tr>
<tr>
<td>The Purpose</td>
<td>6</td>
</tr>
<tr>
<td>Justification for the Study</td>
<td>6</td>
</tr>
<tr>
<td>Definitions</td>
<td>7</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>9</td>
</tr>
<tr>
<td>Assumptions</td>
<td>10</td>
</tr>
<tr>
<td>Delimitations</td>
<td>11</td>
</tr>
<tr>
<td>Summary</td>
<td>11</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>13</td>
</tr>
<tr>
<td>Introduction</td>
<td>13</td>
</tr>
<tr>
<td>Literature Related to Reading Comprehension Skills</td>
<td>13</td>
</tr>
<tr>
<td>Literature Related to Objectives in Education</td>
<td>20</td>
</tr>
<tr>
<td>Literature Related to Criterion-Referenced</td>
<td>44</td>
</tr>
<tr>
<td>Measurement in Education</td>
<td>61</td>
</tr>
<tr>
<td>Summary</td>
<td>61</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>63</td>
</tr>
<tr>
<td>Introduction</td>
<td>63</td>
</tr>
<tr>
<td>Population and Sample</td>
<td>64</td>
</tr>
<tr>
<td>Experimental and Control Group Procedures</td>
<td>71</td>
</tr>
<tr>
<td>Research Design</td>
<td>77</td>
</tr>
<tr>
<td>Sources of Data</td>
<td>79</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>81</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>84</td>
</tr>
<tr>
<td>Statistical Analysis of Data</td>
<td>85</td>
</tr>
<tr>
<td>Summary</td>
<td>86</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>SES Index of Participating Schools, 1976-77</td>
<td>68</td>
</tr>
<tr>
<td>2.</td>
<td>Number of Subjects Classified by Grade, Reading Approach, and Socioeconomic Level</td>
<td>69</td>
</tr>
<tr>
<td>3.</td>
<td>Distribution of Subjects</td>
<td>70</td>
</tr>
<tr>
<td>4.</td>
<td>CRT Means for the Sample Classified by Approach by Grade Level and Approach by SES</td>
<td>91</td>
</tr>
<tr>
<td>5.</td>
<td>Summary Table of the Analysis of Covariance of CRT Data for Approach, Grade Level, and SES</td>
<td>92</td>
</tr>
<tr>
<td>6.</td>
<td>CRT Adjusted Means for the Experimental and Control Groups</td>
<td>93</td>
</tr>
<tr>
<td>7.</td>
<td>CRT Adjusted Means for Grades Four and Five</td>
<td>94</td>
</tr>
<tr>
<td>8.</td>
<td>CRT Adjusted Means for Low and Middle SES</td>
<td>95</td>
</tr>
<tr>
<td>9.</td>
<td>CRT Gain Score Means Classified by Approach and Grade Level</td>
<td>96</td>
</tr>
<tr>
<td>10.</td>
<td>CRT Gain Score Means Classified by Approach and SES</td>
<td>97</td>
</tr>
<tr>
<td>11.</td>
<td>NRT Means for the Sample Classified by Approach by Grade Level and Approach by SES</td>
<td>99</td>
</tr>
<tr>
<td>12.</td>
<td>Summary Table of the Analysis of Covariance of the NRT Data for Approach, Grade Level, and SES</td>
<td>100</td>
</tr>
<tr>
<td>13.</td>
<td>NRT Adjusted Means for the Experimental and the Control Group</td>
<td>101</td>
</tr>
<tr>
<td>14.</td>
<td>NRT Adjusted Means for Low and Middle SES</td>
<td>102</td>
</tr>
<tr>
<td>15.</td>
<td>Correlation Matrix for Pretest and Posttest Between the CRT and NRT for the Total Group</td>
<td>103</td>
</tr>
<tr>
<td>16.</td>
<td>Correlation Matrix for Pretest and Posttest Between the CRT and NRT for the Experimental Group</td>
<td>104</td>
</tr>
<tr>
<td>17.</td>
<td>Correlation Matrix for Pretest and Posttest Between the CRT and NRT for the Control Group</td>
<td>105</td>
</tr>
</tbody>
</table>
LIST OF FIGURES AND CHARTS

<table>
<thead>
<tr>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Graphic Representation of the Approach by Grade Interaction on the CRT</td>
<td>96</td>
</tr>
<tr>
<td>2. Graphic Representation of the Approach by SES Interaction on the CRT</td>
<td>98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chart</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparison of Comprehension Models</td>
<td>18</td>
</tr>
<tr>
<td>2. Summary Chart of the Hypotheses, Findings, and Conclusions</td>
<td>114</td>
</tr>
</tbody>
</table>
Since the turn of this century, the educational community has witnessed many and varied efforts to improve instruction which have resulted in several notable developments. In reading, for example, there has been a proliferation of teaching approaches introduced and developed since 1910. Smith traced these developments, decade by decade: between 1910 and 1920, the teaching of silent reading was introduced; between 1920 and 1930, the emphasis on individual differences resulted in instructional grouping for reading and the introduction of remedial reading instruction; between 1930 and 1940, the reading readiness concept was widely accepted and implemented; between 1940 and 1950, reading began to be considered as a part of the language arts.

An examination of the next two decades reveals additional developments in reading, among them individualized reading, reading programs including different ethnic groups, linguistic reading programs, programmed reading, reading programs using Initial Teaching Alphabet, and the language experience approach to teaching reading.

---

1Nila B. Smith, American Reading Instruction (Newark, Del.: International Reading Association, 1965), pp. 157-415.
Teaching by means of the objectives-based procedure received considerable emphasis from the early 1960's as curriculum and instruction leaders were seeking to clearly define instruction in terms of resulting student behavior.\(^2\) Hambleton elaborated on this concept:

The overall goal of an objectives-based instructional program is to provide an educational program which is maximally adaptive to the requirements of the individual learner. The instructional objectives specify the curriculum and serve as a basis for the development of curriculum materials and achievement tests.\(^3\)

Emphases on objectives-based instruction stimulated interest in the concept of criterion-referenced measurement as specialists sought additional approaches to the measurement of learning outcomes.\(^4\)

Hambleton stated that

... one of the underlying premises of objectives-based programs is that effective instruction depends, in part, on a knowledge of what skills the student has. It follows that the tests used to monitor student progress should be closely matched to the instruction.\(^5\)

A number of basic influences were instrumental in focusing attention on reading instruction as specified by Smith: (1) expanding


knowledge, (2) technological revolution, (3) national concern—the welfare of the country in relation to other powers, (4) pressures following the flight of Sputnik in 1957, and (5) the impetus of governmental support of education for the masses.  

These influences may also have been instrumental in focusing attention on objectives-based instruction and criterion-referenced measurement. Specifically related to governmental support of education was the enactment of the Elementary and Secondary Education Act in 1965, with its unequivocal emphasis on cognition in reading and mathematics in compensatory education schools. This legislation not only brought about more intensive emphases on the improvement of reading instruction, but served to focus additional attention on and gain support for implementation of curricula through the use of instructional objectives.

In California, more support for objectives-based instruction came primarily from two unrelated events: (1) the passage of the Stull Bill in 1968, which called for local school districts to develop guidelines by which the performance of certificated personnel would be evaluated in accordance with the degree of student attainment of objectives; and (2) the call for innovative approaches to teaching in compensatory education schools by the Division of Compensatory Education.
California State Department of Education, which would focus on the needs of low-achieving children in reading and mathematics.  

Many educators responded to these influences and events by developing objectives-based instructional systems in reading and/or other subject areas. Three of the most well known and widely used are: (1) Individually Prescribed Instruction (IPI), which was developed at the Learning Research and Development Center at the University of Pittsburgh, and initiated during the early 1960's at Oakleaf School in cooperation with the Baldwin-Whitehall Public School District near Pittsburgh; (2) Program for Learning in Accordance with Needs (PLAN), which was developed in the 1960's by the American Institutes for Research, the Westinghouse Learning Corporation, and twelve school districts in Northern California; and (3) Mastery Learning, which was introduced in the 1920's in the format of the Winnetka Plan in Winnetka, Illinois. This last system did not become objectives-based until the late 1960's. 

The present research was concerned with the effectiveness of the Project: Individualized Reading and Mathematics Interdistrict (PIRAMID) objectives-based instructional system. The system was developed by teachers in a consortium of seven school districts, and implemented in

---


11These were the Bakersfield City School District, the Compton Unified School District, the National School District, the Berkeley
fall of 1972. Distribution of materials and other information regarding PIRAMID is managed by the Consortium Office, located in Yuba City, California.

The present quasi-experimental study was designed to compare the effects on reading achievement of two approaches to teaching reading comprehension—the PIRAMID objectives-based approach, and a non-objectives-based approach—and to determine the extent to which a Criterion-Referenced Test (CRT) and a Norm-Referenced Test (NRT) agree in assessment of reading comprehension achievement. The PIRAMID instructional objectives and criterion-referenced tests are contained in Appendix A. The CRT answer sheets will be found in Appendix B.

THE PROBLEM

The problem was divided into two parts. The first part investigated the question whether students who were taught reading comprehension skills by means of instructional objectives showed a greater gain than students who were taught the same skills by means of suggestions from a basal reader teacher's guide. Those tested were fourth and fifth grade students of middle and low socioeconomic status. The second part was to ascertain the degree of concordance between results of the criterion-referenced measure and results of the norm-referenced measure.

Unified School District, Kern County Cooperatives, the Santa Ana Unified School District, and the Yuba City Unified School District, all in California.
THE PURPOSE

The purpose was twofold: (1) to compare gains in reading comprehension achievement of an experimental group with gains of a control group, and (2) to compare the pretest and posttest scores of the CRT with those of the NRT to determine the extent of relationship between the two measures. The experimental group used the PIRAMID objectives and the control group used suggestions from a basal reader teacher's guide.

JUSTIFICATION FOR THE STUDY

The PIRAMID instructional system was adopted by sixty-six school districts in California, Arizona, Colorado, Louisiana, and the American schools in Singapore in the 1976-77 school year. The value of this objectives-based system to teaching and learning proposed an empirical investigation since no such evidence existed.

For the edification of these school districts, the PIRAMID Consortium, and the larger population, the present study was an attempt to provide answers to the following research questions:

1. Does teaching reading comprehension by the objectives-based procedure enhance the performance of students in the fourth and fifth grades?

2. Does teaching reading comprehension by the objectives-based procedure equally enhance the performance of students belonging to middle and low socioeconomic groups?

3. To what extent do the PIRAMID CRT and the NRT agree in the assessment of reading achievement?
DEFINITIONS

The following terms were defined as they were used in this study:

**Basal Reading.** Basal reading was defined as a plan for teaching the skills of developmental reading through the use of a series of graded textbooks, workbooks, and planned activities.\(^\text{12}\)

**Compensatory Education.** Compensatory education was defined as:

... programs that focus on the educational needs of pupils who are potentially able to succeed in school but who, because of lingual, cultural, economic, and environmental handicaps, are unlikely to succeed without special programs.\(^\text{13}\)

**Comprehension.** Reading comprehension encompassed a three-part definition: literal reading, interpretive reading, and critical reading. Each paralleled the classification and structure of the PIRAMID reading comprehension objectives. A definition of each part follows:

1. **Literal reading** refers to the acquisition of stated ideas and information. It includes word meaning, contextual clues, sentence meaning, and paragraph organization.

2. **Interpretive reading** involves implied meanings or reading "between and beyond the lines." It involves reading for inferences—drawing conclusions, making generalizations, recognizing the author's purpose, and anticipating outcomes.

3. **Critical reading** is the process of examining verbal materials in the light of related objective evidence, comparing the statement with some norm or standard, and concluding or acting upon the judgment then made. It involves judging accuracy, recognizing facts and opinions, and recognizing persuasive statements.\(^\text{14}\)

---


\(^{13}\) Office of Compensatory Education, Highlights, p. 1.

Criterion-Referenced Test (CRT). A criterion-referenced test was defined as "a measure which is used to diagnose what an individual can or cannot do with respect to some established criterion rather than with respect to other individuals." 15

Developmental Reading. Developmental reading was defined as:

... a program in which students who are able readers continue to be taught reading skills in a sequential program of instruction, designed to reinforce and extend the skills and appreciations acquired in previous years, and to develop new skills as they are needed. It should emphasize the development of reading power and guide students in the selection of reading materials. 16

Instructional Objective. An instructional objective was defined as "an objective that specifies under what condition and to what extent a certain kind of student performance can be expected to take place." 17

Instructional System. An instructional system consists of a skill continuum of instructional objectives, suggestions for prescribing learning activities for student attainment of the objectives, suggested class management procedures, and CRT's.

Norm-Referenced Test (NRT). A norm-referenced test was defined as an instrument which is used to ascertain an individual's performance in relation to the performance of other individuals on the same measuring device. The meaningfulness of the individual score emerges

---


from the comparison. It is because the individual is compared with some normative groups that such measures are described as norm-referenced. Most standardized tests of achievement or intellectual ability can be classified as norm-referenced measures.\textsuperscript{18}

Objectives-Based Instruction. Objectives-based instruction was defined as instruction specified "in terms of instructional objectives . . . ."\textsuperscript{19}

Socioeconomic Status (SES). Socioeconomic status was defined as the social and economic level of the subjects, as determined by the occupation of their parents or guardians. The method used for determining SES was the occupation classification system developed by the California State Department of Education.\textsuperscript{20} A discussion of this method of determining socioeconomic status is contained in Chapter III.

**HYPOTHESES**

The following hypotheses were developed to guide the investigation of the problem and to accomplish the purposes of the study.

**Hypothesis 1.** The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the PIRAMID CRT.

**Hypothesis 2.** There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when

\textsuperscript{18} Popham and Husek, "Implications of Criterion-Referenced Measurement, p. 2.

\textsuperscript{19} Hambleton and others, Criterion-Referenced Testing, p. 10.

Hypothesis 3. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the CRT.

Hypothesis 4. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the reading comprehension subtest of the NRT.

Hypothesis 5. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the NRT.

Hypothesis 6. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the NRT.

Hypothesis 7. There is no correlation between pretest and post-test measures on reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT.

Hypothesis 8. The correlation between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT is equivalent for the experimental group and the control group.

ASSUMPTIONS

The following four assumptions were relevant to this investigation: (1) The student sample was representative of the target population from which it was selected. (2) The sample teachers and
schools were representative of target schools and teachers from which they were selected. (3) The pretest sensitization and "Hawthorne Effect" were minimal. (4) The norming samples for both NRT's were selected from similar populations.

DELIMITATIONS

Primarily, the generalization values of the findings and conclusions are delimited by the population selected for the study but may offer useful information for a larger population. The sample of 670 fourth and fifth grade students attended eight schools: three in the Hayward Unified School District, Alameda County; four in the Pittsburg Unified School District, Contra Costa County; and one in the Yuba City Unified School District, Sutter County, all in California.

The extent of cooperation and voluntarism of participating teachers is not known. Therefore, these additional factors impose delimitations on the generalizability of the findings and conclusions.

SUMMARY

The first chapter served as an introduction to the study; it provided a statement of the problem, a statement of the purpose, justification for the study, definitions of terms used, a statement of the hypotheses, assumptions of the study, and delimitations. Chapter II consists of a review of related literature which includes (1) literature related to reading comprehension skills, (2) literature related to objectives in education and a review of previous studies which have investigated the effects of instructional objectives on learning, and (3) literature related to criterion-referenced measurement in education.
and a review of previous studies which have investigated the relationship between CRT's and NRT's. Chapter III contains a discussion of methodology, which includes a restatement of the problem and purpose, a discussion of the population and sample, experimental and control group procedures, a discussion of the research design, sources of data, a description of instruments used, a restatement of the hypotheses, and statistical analysis of data. Chapter IV reports the findings of the investigation related to the stated hypotheses; and Chapter V includes the conclusions, administrative implications, and recommendations.
CHAPTER II

REVIEW OF RELATED LITERATURE

INTRODUCTION

The review of related literature is divided into three parts: (1) literature related to reading comprehension skills; (2) literature related to objectives in education and a review of previous studies which have investigated the effects of instructional objectives on learning; and (3) literature related to criterion-referenced measurement in education and a review of previous studies which have investigated the relationship between CRT's and NRT's.

LITERATURE RELATED TO READING COMPREHENSION SKILLS

The views of several reading authorities were considered for the purpose of determining what constitutes reading comprehension. Their opinions were diversified. Their judgments were useful in determining the types and complexity of skills included and for disclosing those views which were in conformity with the PIRAMID comprehension objectives.

Reading comprehension is a two-level process: objective comprehension and subjective comprehension,¹ a process of understanding

---

and interpreting, literal and interpretive.

Others describe reading comprehension as a three-level process.

The definition presented by Karlin is restated here:

1. **Literal reading** refers to the acquisition of stated ideas and information. It includes word meaning, contextual clues, sentence meaning, and paragraph organization.

2. **Interpretive reading** involves implied meanings or reading "between and beyond the lines." It involves reading for inferences—drawing conclusions, making generalizations, recognizing the author's purpose, and anticipating outcomes.

3. **Critical reading** is the process of examining verbal materials in the light of related objective evidence, comparing the statement with some norm or standard, and concluding or acting upon the judgment then made. It involves judging accuracy, recognizing facts and opinions, and recognizing persuasive statements.

Other proponents of the three-level process defined comprehension as literal, interpretive, and evaluation, or literal, interpretation, and problem solving. Several gave more expanded lists of skills considered to be basic to understanding.

---


6 Robert M. Wilson, *Diagnostic and Remedial Reading for Classroom and Clinic* (Columbus, Ohio: Charles E. Merrill, 1972), pp. 234-43.

All of the aforementioned were concerned with reading remediation with the exception of Zintz, Karlin, and Harris and Sipay, whose definitions are applicable to children in a developmental reading program.

Several concerned with developmental reading consider comprehension to be a thinking process and a problem-solving activity, but in substance their explanations include the literal, interpretive, and critical reading skills.\(^8\)

A different notion was presented by Spache and Spache,\(^9\) who defined reading comprehension as comprehension and critical reading, suggesting that critical reading is a type or degree of comprehension.

Although it has not been established that reading comprehension can be broken down into discrete levels or factors,\(^10\) most authorities consider their level differentiation as more a continuum than as

---


discrete levels. It is interesting also that Davis, who conducted a study of the skills basic to the ability to comprehend, found the most significant factors to be knowledge of word meanings and ability to draw inferences.

Several reading series were examined for comprehension skills considered important to their program. Several used detailed many-faceted definitions, while others emphasized the three-level approaches.

In addition to examining definitions of reading comprehension as given by various reading authorities and publishers, several comprehension models were noted for comparison.

A committee of educators headed by Bloom developed a model of six major classes of objectives, each at increasing levels of complexity.

11 Frederick B. Davis, "Research in Comprehension in Reading," Reading Research Quarterly, III (Summer, 1968), 499-545.

12 David H. Russell and others, The Ginn Basic Readers (Boston: Ginn & Co., 1966); Margaret Early and others, The Bookmark Reading Program (San Francisco: Harcourt, Brace, Jovanovich, 1974); Irma S. Black and others, eds., The Bank Street Reading Series (New York: Macmillan, 1966); and Ida Mae Johnson and others, The New Open Highways (Oakland, Calif.: Scott, Foresman, 1974).


Beyond the recognition of words, the model can be applied directly to reading comprehension. The model is contained in Appendix C.

A Taxonomy of the Cognitive and Affective Dimensions of Reading Comprehension was developed by Barrett. The model consists of five levels increasing in difficulty from literal comprehension to appreciation. In developing the model, Barrett drew heavily on Bloom's ideas. Barrett's model will be found in Appendix D.

Guilford classified the factors of intelligence according to the basic kind of process or operation performed. The model suggests that content is dealt with on a unit, or larger, basis, and has five main processes or operations. The processes are listed in Appendix E.

These were noted because Karlin has stated that "reading models and research on reading comprehension provide some clues as to what factors and processes might be operating and accounting for variations in reading performances." 

The models developed by Bloom, Barrett, and Guilford are shown in Chart 1, and are compared with Karlin's definition of reading comprehension. Note that Karlin's definition contains most of the factors and processes shown in the models. Note also that Barrett's affective category, appreciation, is not incorporated in any of the other schemes.


17 Karlin, Teaching Elementary Reading, p. 182.
<table>
<thead>
<tr>
<th>Karlin's Definition of Reading Comprehension</th>
<th>Bloom's Taxonomy of Objectives</th>
<th>Barrett's Taxonomy of Affective Dimensions</th>
<th>Guilford's Structure of Intellect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4. Analysis</td>
<td>3. Inferential comprehension</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. Synthesis</td>
<td></td>
<td>4. Divergent thinking</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and therefore this category is not included in the current study.

Karlin also stated that "unless reading comprehension is described in operational terms, it is not possible to plan instruction to increase comprehension nor determine how well children do comprehend what they read," suggesting that the factors spelled out in the models are essential in defining reading comprehension if reading comprehension abilities are to be increased.

Summary

In the preceding section, definitions of reading comprehension were discussed. Definitions presented were: (1) those most appropriate for remedial and/or corrective reading; (2) those most applicable to children in a developmental reading program; (3) those emphasized by several publishers of basal reading series; and (4) several reading comprehension models. While there is little agreement among reading authorities on what constitutes reading comprehension, there is considerable agreement among some authorities and the developers of reading comprehension models. The latter classify skills according to increasing levels of complexity. The definition of reading comprehension used in the current study was compared with the reading comprehension models. This was done to note that this definition deals with not only the simpler literal behaviors but also with the more complex interpretive and critical behaviors. It was stated that Barrett's appreciation category was not included in the other schemes and hence was not included in the current study.

---

18 Karlin, Teaching Elementary Reading, p. 182.
Probably the educational community has always realized the importance of having a sense of purpose since general elements of purpose have been developed and used extensively. One of the earliest and most important of these was introduced in 1918 by the Commission on the Reorganization of Secondary Education of the National Education Association. The Commission stated the following objectives of secondary education, which have come to be known as The Cardinal Principles of Education:

1. Health
2. Command of the fundamental processes
3. Worthy home membership
4. Vocation
5. Civic education
6. Worthy use of leisure
7. Ethical character

A number of other sets of educational objectives have been prepared. In 1924, for example, Bobbitt developed a tenfold classification:

1. Language activities
2. Health activities
3. Citizenship activities
4. General social activities
5. Spare-time activities
6. Keeping one's self mentally fit
7. Religious activities
8. Parental activities, the upbringing of children, the maintenance of a proper home life
9. Unspecialized or nonvocational practical activities
10. The labors of one's calling

---


20 Franklin Bobbitt, How To Make a Curriculum (San Francisco: Houghton Mifflin, 1924), pp. 8-9.
The Educational Policies Commission of the National Education Association published a list of objectives for education fourteen years later, listed under four main categories, as follows:

1. Self-realization
2. Human relationships
3. Economic efficiency
4. Civic responsibility

The Imperative Needs of Youth appeared in a later document of the Educational Policies Commission. Although definitely for the secondary school, they are also of importance for the elementary school:

1. All youth need to develop salable skills and those understandings and attitudes that make the worker an intelligent participant in economic life.

2. All youth need to develop and maintain good health and physical fitness.

3. All youth need to understand the rights and duties of the citizen of a democratic society, and to be diligent and competent in the performance of their obligations as members of the community and citizens of the state and nation.

4. All youth need to understand the significance of the family for the individual and society and the conditions conducive to successful family life.

5. All youth need to know how to purchase and use goods and services intelligently, understanding both the values received by the consumer and the economic consequences of their acts.

6. All youth need to understand the methods of science, the influence of science on human life, and the main scientific facts concerning the nature of the world and of man.

7. All youth need opportunities to develop their capacities to appreciate beauty in literature, art, music, and nature.

---


8. All youth need to be able to use their leisure time well and to budget it wisely, balancing activities that yield satisfactions to the individual with those that are socially useful.

9. All youth need to develop respect for other persons, to grow in their insight into ethical values and principles, and to be able to live and work cooperatively with others.

10. All youth need to grow in their ability to think rationally, to express their thoughts clearly, and to read and listen with understanding.

Taba advanced the notion that "in order for objectives to serve their functions well, a systematic approach to the formulation and organization of objectives is needed." She recommended the following set of standards useful in formulating educational objectives:

1. A statement of objectives should describe both the kind of behavior expected and the content or the context to which that behavior applies.

2. Complex objectives need to be stated analytically and specifically enough so that there is no doubt as to the kind of behavior expected, or what the behavior applies to.

3. Objectives should be so formulated that there are clear distinctions among learning experiences required to attain different behaviors.

4. Objectives are developmental, representing roads to travel rather than terminal points.

5. Objectives should be realistic and should include only what can be translated into curriculum and classroom experience.

6. The scope of objectives should be broad enough to encompass all types of outcomes for which the school is responsible.

---


24 Ibid., pp. 200-5.
According to Duchastel and Merrill, emphasis on the specific and more clearly stated objective began with the publication in 1962 of Mager's book on Preparing Instructional Objectives. Bobbitt, however, had expressed this need as early as the 1920's. Further, Kliebard stated that "the first issue of the Review of Educational Research in 1931 . . . clearly subscribed to the notion that educational objectives must be specific, detailed, and determine in advance of all other curriculum planning." Later, Smith and Tyler advanced the notion that since the educational process seeks to change the behavior patterns of human beings, the kinds of changes in behavior patterns ought to be reflected in its educational objectives.

According to Kliebard, Tyler in the early 1950's issued a strong plea that curriculum planning begin with a statement of educational objectives in behavioral terms. This was done along with a list of steps Tyler used in curriculum planning, which came to be known as the Tyler rationale.


26 Bobbitt, How To Make a Curriculum, p. 243.


29 Kliebard, "Curricular Objectives and Evaluation," p. 244.
Support for Tyler's plea came from Bloom; Gagné, who devoted considerable attention to describing a "task analysis" of what a student is expected to do when he has successfully completed a unit of instruction; Krathwohl and Masia, who, along with Bloom, developed a model of five major classes of objectives useful in evaluating objectives in the affective domain; Mager and Popham, who developed guidelines for writing instructional objectives; and Kibler, Barker, and Miles, who developed levels of objective specificity.

Popham discussed the terms used to describe instructional objectives. He said that

the most knowledgeable proponents of explicit instructional objectives have veered away from using the phrase behavioral objectives, for they recognize that some educators erroneously equate the adjective "behavioral" with a mechanistic, dehumanized form of behaviorism. . . . Thus, because such phrases create less misdirected resistance, expressions similar to

---

30 Bloom, ed., Taxonomy of Educational Objectives, p. 5.
"performance objectives," "measurable objectives," or "operational objectives" are often employed these days.36

The term "instructional objectives" was chosen for use in the present study to avoid the resistance mentioned by Popham. The term was chosen also because it is widely used in the literature to describe the explicitly stated objective.

Even though there is strong support for using instructional objectives, a small group of educators have expressed reservations about their value to teaching and learning.37 Recently, Durio advanced the notion that "there is more emphasis on detail than underlying structure, with excessive attention to using specific terms, rather than providing for transferral of the knowledge beyond one instructional segment,"38 and Campbell succinctly summed up his reservation, stating that instructional objectives disregard contemporary research which indicates that effective learning requires a highly individualized and flexible mode of instruction. Insistence on strict behavioral objectives makes it difficult for teachers to be flexible or indeed to truly meet the needs of individual students.39


38 Helen F. Durio, "Behavioral Objectives: Where Have They Taken Us?" The Clearinghouse, XLIX (January, 1976), 201.

Nevertheless, Eisner has suggested that whether instructional objectives are of value or not to teaching and learning is an empirical question. Therefore, numerous investigators have turned to research in an attempt to base their perceptions of the issue on empirical grounds.

Studies Related to the Effect of Objectives on Learning

In an extensive review of research studies which addressed the issue of effects of instructional objectives on learning, Duchastel and Merrill examined many studies. Ten studies investigated the hypothesis that students provided with objectives would achieve more than students not provided with them; seven investigated interactions between type of learning and availability of objectives; eight investigated interactions between availability of objectives and certain learner characteristics; and three investigated the hypothesis that students provided with objectives will take less time to learn instructional material than students without objectives. There was inconsistency in the findings of the various studies.

Studies which found no significant differences between the experimental and the control groups are as numerous as those which have found such a difference. Furthermore, when we consider the total number of studies which have investigated effects on student achievement, an even smaller proportion of studies have found a significant main effect for this variable. However, those studies which have found such an effect have usually favored the presentation of objectives.

---

42 Ibid., p. 63.
Duchastel and Merrill suggested that "future research should focus attention on student use of objectives as these objectives will make no difference in achievement if students pay no attention to them."\(^{43}\)

Subsequent to the Duchastel and Merrill review, several researchers investigated the hypothesis that achievement of students will increase if they make use of instructional objectives during the learning situation. Therefore, they favored the presentation of objectives.

In a study involving graduate students who were enrolled in a course required for teacher certification, the students were asked to rank their acceptance of each of thirty objectives by placing six in each of five categories ranging from most acceptable to least acceptable. The students were asked to perform this task at the beginning and at the end of each class session. Test scores for each of the five groups of objectives were individually ranked and summarized, using the Kendall Coefficients of Concordance, Form W, corrected for ties in ranks. It was concluded that there was no significant difference at the .05 level of significance between student ranking of instructional objectives as to their acceptance at the beginning or at the termination of the course and their consequent scores on the final examination or test of recall one month after the final exam.\(^{44}\)

\(^{43}\) Duchastel and Merrill, "The Effects of Behavioral Objectives on Learning," p. 65.

\(^{44}\) Terry L. Gibson, "Effect upon Learning of Student Knowledge and Acceptance of Behavioral Objectives," (paper presented at the Association for Educational Communications and Technology Annual Convention, Atlantic City, 1974).
Duell provided instructional objectives to 167 college seniors who participated in two separate experiments. Experiment One was designed to investigate the joint effects of the level of the test questions and the availability of instructional objectives on learning from written prose. Fifty-six students enrolled in two sections of an educational psychology course taken just before student teaching were randomly assigned to two experimental conditions: an instructional objectives group, and a noninstructional objectives group. Each experimental condition contained twenty-eight subjects. An additional group of twenty-four students in another section of the same course acted as a control group. The prediction was that the difference between the instructional objectives and noninstructional objectives groups would be nonsignificant for recall questions but significant for application questions. The two experimental groups combined did perform significantly better than the control group on the posttest. There was a nonsignificant difference between the instructional and noninstructional objectives groups on the application questions, and a significant difference on the recall questions, with the instructional objectives group performing significantly better than the noninstructional objectives group.

Experiment Two was designed to test the hypothesis that the judged importance of an item of information determines whether knowledge of instructional objectives during training is helpful. Thirty college seniors were randomly assigned to the instructional objectives

condition, thirty were randomly assigned to the noninstructional objectives group, and twenty-seven constituted the control group. On test data, the two experimental groups performed significantly better. The instructional objectives group performed significantly better than the noninstructional objectives group on the unimportant recognition questions, while the differences between these two groups on the important recognition questions and the important application questions were nonsignificant. Posttest-delayed-posttest comparisons revealed significant losses for the instructional objectives and noninstructional objectives groups combined on the unimportant recognition questions and the important recognition questions, but no significant loss on the important application questions where a slight gain was shown due to the instructional objectives group. It was concluded that the data of Duell’s study support the hypothesis that the judged importance of an item of information determines whether knowledge of instructional objectives during training is helpful. Students receiving instructional objectives during study performed significantly better on test questions that the majority of the students classified as unimportant than students receiving noninstructional objectives.

Merrill Investigated the effects of presentation of instructional objectives on the learning process. One hundred thirty subjects were taken from four sections of an introductory educational psychology course at the University of Texas at Austin. They were administered six ability tests and randomly assigned to an example-only, an

---

objective-example, a rule-example, or an objective-rule-example treatment. In addition, data were obtained for each subject on the following criteria: total number of examples required to learn the science and display latency, test-item response latency, and total latency. Display latency was the total time the subject spent studying the examples. Test-item response latency was the total time required by the subject to respond to the three-item tests following each example display. Total latency was merely the sum of the display and the test-item response latencies. Rules significantly reduced the number of examples and the total time required to complete the task and increased performance on a transfer test. Instructional objectives did not affect significantly total or display latency, but significantly reduced test-item response latency and the required number of examples.

Kaplan\(^{47}\) examined the effects of part versus whole presentations of instructional objectives and text upon intentional and incidental prose learning. The sample consisted of 540 paid volunteers, between fifteen and eighteen years of age, from six New Jersey schools. Eighteen subjects were assigned to each of twenty-four treatments (\(N=432\)). In addition, 108 subjects served in six control groups who read the passages without instructional objectives. The 2 by 3 by 2 by 2 analyses of variance were performed with two levels of presentation (part and whole), three levels of passage length (56, 113, and 169 sentences), two levels of objective specificity (specific and general),

and two levels of density (40% and 60%). Separate analyses were performed for intentional learning, incidental learning, and inspection time. Results of analysis of data on intentional learning showed that part presentation resulted in greater learning than whole presentation. The main effect of passage length was highly significant. Results of analysis of data on incidental learning showed that there was no significant difference for incidental learning between part and whole presentations. The passage length main effect was significant. Results of analysis of data on inspection time showed that passages and objectives with part presentations required more reading time than with whole presentations. Shorter passages required less reading time than longer passages. General objectives required less time than instructional objectives. Density 40 percent required less time than Density 60 percent. None of the interactions was significant.

Daniel designed a study to determine the extent to which knowledge of instructional objectives influences learning outcomes in the cognitive and affective areas in a religious educational setting. Three groups of adults attending Bible classes were taught under three different conditions for six weeks: (1) neither teacher nor students had knowledge of the objectives; (2) both teacher and students received two sets of general objectives (one cognitive, the other affective) prior to instruction; and (3) both teacher and students received two sets of general and instructional objectives (one cognitive,

---

the other affective) prior to instruction. A pretest and a posttest were
given to each group. No differences were found on cognitive outcomes due
to knowledge of instructional objectives, neither were any differences
found for the affective objective of awareness. However, significant
differences due to knowledge of instructional objectives were found for
the affective objectives of appreciation, satisfaction, and devotion.
Significant differences existed in each of three cases between the group
taught with knowledge of general and instructional objectives and the
group which received only general objectives. Significant differences
also existed between the group which received general and instructional
objectives and the group with no knowledge of objectives for the affec-
tive objectives of satisfaction and devotion.

Rosen provided students with instructional objectives and
reinforcement for correct test responses in the form of monetary compen-
sation and early class dismissal. He also provided practice items for
some students in order to demonstrate the relevancy of instructional
objectives to test performance. Instructional objectives did not have
an effect on the amount learned. Reinforcement actually decreased
learning.

Petty examined the differences in student academic achievement
and student attitude in an instructional media course between those who

49 Theodore A. Rosen, "The Effect of Instructional Objectives,
Reinforcement, and Test Items on Learning from Text" (unpublished

50 Bruce A. Petty, "An Investigation of the Effects of Written
Behavioral Objectives upon Performance and Attitudes of Students in an
Instructional Technology Course at Kansas State University," Dissertation
Abstracts, XXXVI (July, 1975), 149-50-A.
were given written instructional objectives and those who were not given
the objectives. The data were also used to test the existence of a
correlation between student attitude and student performance, and to
determine its magnitude if such a correlation existed. A posttest-only
control group design was used. The t-test for independent samples was
used to measure differences in criterion means, the Pearson Product-
Moment statistic to determine correlations, and a Coefficient Alpha to
determine reliability of the instruments used. The findings indicated
no significant difference in attitude survey means between experimental
and control groups. There was a significant correlation between
student attitude and student performance for those students who received
written instructional objectives and for those who did not receive
them.

Gordon\textsuperscript{51} designed a study to determine the extent to which
teacher-prepared instructional objectives, provided to students prior
to studying written assignments, affect initial acquisition and retention
of students having varying abilities. Forty-eight seventh grade indus-
trial arts students were classified high mental or low mental ability
and randomly assigned to treatment either with or without instructional
objectives. Students were subjected to the treatment for three
consecutive days, each day culminating with a test of initial acquisi-
tion. Eight days following the treatment, a test of retention was
administered. Mean initial acquisition and retention scores were

\textsuperscript{51}Robert A. Gordon, "Effects of Instructional Objectives on
Reading Acquisition and Retention Levels of Students with Varying
analyzed, using analysis of variance. Students with high mental ability scored significantly higher on both initial acquisition and retention tests than students with low mental ability whether they were exposed to the objectives or not. No significant differences resulted between mean scores of high or low mental ability groups with and without exposure to objectives.

Song investigated the effects of presentation of instructional objectives, introduction to films, pretest, and programmed instruction text prior to actually showing the instructional film in learning American history. The study included three separate experiments:

1. The first experiment involved three different instructional films with three variables, i.e., instructional objectives, introduction, and pretest. The sample consisted of 183 eighth graders who were randomly assigned to four groups. Results of the first experiment showed that the effectiveness of instructional objectives, introduction, and pretest vary according to each different film. An analysis of variance on the data of Film B revealed significant effects of previously mentioned variables on the posttest; however, there were no significant differences among the variables in the results of Film A and Film C on their respective posttests.

2. The second experiment used 212 seventh graders with Film A. An analysis of variance showed that there were no significant differences between the performance of the four groups.

3. The third experiment involved a linear programmed instruction text which revealed significant effects on learning from Film C. However, the effect of the second variable, instructional objectives, was not significant. The 185 seventh graders in the sample were tested with a pretest; based on the results, students were categorized as high and low prior knowledge groups. Students in the high prior knowledge group performed significantly better than those in the low prior

knowledge group on the posttest; however, there were no treatment/aptitude interactions among all students, regardless of their group category.

Kaplan and Simmons\(^{53}\) investigated the learning effects of instructional objectives presented prior to or after a text when the objectives were written with or without relevant information. Four experimental treatments were examined: objectives presented before text (a) with and (b) without relevant information and objectives presented after text (c) with and (d) without relevant information. Subjects were tenth, eleventh, and twelfth grade students from three New Jersey high schools. A total of three hundred students were randomly assigned to all treatments. This allowed for sixty subjects in each of the four treatment groups and the control group. Results of the 2 by 2 by 2 analysis of variance showed no differences in performance (a) between objectives located prior to and after the passage, or (b) between objectives containing information and those not containing information. However, there was a significant location of objectives by information content interaction. The main finding was that performance on information relevant to an objective was relatively high whether the objectives were presented before or after the text. Performance on incidental material was greater for objectives located after the text than before the text. More inspection time was consumed by the experimental groups than by a group that received no objectives.

\(^{53}\) Robert Kaplan and Francine G. Simmons, "Effects of Instructional Objectives Used as Orienting Stimuli or as Summary/Review upon Prose Learning," *Journal of Educational Psychology*, LXVI (August, 1974), 614-22.
Nassif\(^{54}\) examined the effects of specificity and position of written instructional objectives on learning from an audiotaped lecture. The sample consisted of 160 undergraduate psychology students who were randomly assigned to one of five experimental groups. Subjects received instructional or general objectives before or after the sections of the lecture. A control group received no objectives. A multivariate analysis of covariance produced a significant effect due to position of objectives favoring the groups receiving objectives before the text. In addition, position resulted in significantly different intentional learning for the treatment groups. Univariate analysis of variance indicated that incidental learning was significantly higher than intentional for the treatment groups combined.

Arp\(^{55}\) examined the effect of providing 138 high school typewriting students with information concerning instructional objectives. A secondary purpose was to compare the more capable typewriting students with those less capable with respect to the effect of having information concerning instructional objectives. The student sample was divided into four control sections and four treatment sections. The four treatment sections were taught by objectives, and the four control classes were the nonobjectives groups. An identical pretest and posttest were given to two control sections and two treatment sections. The remaining

\(^{54}\)Paula M. Nassif, "The Effects of Specificity and Position of Written Instructional Objectives on Learning from a Lecture," Dissertation Abstracts, XXXVI (December, 1975), 3523-A.

four sections took only the posttest. A 3 by 2 by 2 analysis of variance with repeated measures was used to analyze the scores of students who took the pretest and the posttest. A 3 by 2 analysis of variance was used to analyze the scores of students who took only the posttest. Grade point average was used to compare the more capable students with those less capable. An analysis of data indicated no significant differences in typewriting achievement between students who were taught by instructional objectives and students who were not taught by instructional objectives. The typewriting achievement of the more capable students and less capable students was not significantly affected by instructional objectives.

Okoduwa conducted a study to determine whether gains in learning from the use of instructional objectives by a specific rationale are greater than those gains when instructional objectives and their rationale are omitted. The sample consisted of sixty-two teacher education students who entered both the experimental and control treatment variables by using a counterbalancing design. The experiment lasted for six weeks. A pretest was administered the first week and a posttest was administered the last week. A videotaped lecture was administered within the four intervening weeks. Each week's lecture had a different content. Only one of the four weeks' results showed a significant difference at the .05 level of significance in favor of the group that used the experimental treatment. Therefore, the null hypothesis that there will

be no significant difference between gain scores of students in experimental and control groups was retained.

Prior to the Duchastel and Merrill review (1973), Morse and Tillman\textsuperscript{57} suggested a short training period to insure that students understand the meaning of instructional objectives and actually use them while learning. Several researchers have tested this hypothesis.

Bassett and Kibler\textsuperscript{58} trained students to use instructional objectives for a unit of instruction. It was hypothesized that they would score significantly higher on an examination consisting of items matched to the objectives than subjects not trained to use the objectives. The sample consisted of 159 undergraduate students in a human communication theory course at Florida State University. The hypothesis was supported by the data.

Cohen and Hillman\textsuperscript{59} conducted a similar study. The sample consisted of fifty-two students from two universities who had demonstrated mastery of the use of instructional objectives. They were randomly assigned to either ten knowledge level objectives (K) or ten above-knowledge level objectives (A) or to a nonobjectives control

\textsuperscript{57} Jean Morse and Murray Tillman, "Effects on Achievement of Possession of Behavioral Objectives and Training Concerning Their Use" (paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972).

\textsuperscript{58} Ronald E. Bassett and Robert J. Kibler, "Effect of Training in the Use of Behavioral Objectives on Student Achievement," \textit{Journal of Experimental Education}, XLIV (Winter, 1975), 12-16.

group (C). All were tested on the same prose material using knowledge and above knowledge level items correlated with each objective. Results of the 3 by 2 analysis of variance indicated significant main and interaction effects, beyond the .05 level of significance on knowledge level test. A Scheffé analysis of the main effects revealed that the K group outperformed the A group, and the same analysis of interaction showed that one A group performed significantly worse than one K group. A questionnaire revealed no significant differences in learning strategies employed. No data suggested a facilitating effect for above knowledge objectives even though most advocates of the use of objectives generally suggest setting objectives above the knowledge level of Bloom's taxonomy.

Viel exposed learners to a brief instructional period dealing with the importance and use of instructional objectives in an effort to determine whether this activity facilitates learning more than when learners receive objectives but without such instruction. A second purpose was to investigate the effects of several kinds of feedback on learning. The sample consisted of one hundred students in grade four. They were randomly assigned to one of ten experimental groups. Results showed a significant main effect for objectives (p < .001) and for feedback (p < .001).

Chick designed a study to determine the effect on achievement

---


61 David S. Chick, "The Effects on Achievement and Retention of Training Fifth-Grade Children To Use Behavioral Objectives in Self-Instructional Geometry," Dissertation Abstracts, XXXI (December, 1975), 3381-82-A.
and retention of possession of instructional objectives and training in their use. A second purpose was to investigate whether training subjects to use instructional objectives has differential effects upon subjects within lower, middle, or higher intelligence groups is consistent for immediate and delayed posttesting. The sample consisted of 108 fifth grade students who were blocked with respect to intelligence, then randomly assigned to treatment. A 3 by 3 by 2 analysis of variance with repeated measures was used to analyze data. The results indicated no significant difference in achievement on all independent variables. Analysis of variance revealed no significant interactions among the three independent variables.

Note that Duell, in his Experiment Two, found a significant main effect for instructional objectives when students were trained to use objectives they had judged to contain unimportant recognition questions. The differences between the instructional objectives group and the noninstructional objectives group on important recognition items and the important application questions were nonsignificant. Both groups, however, performed significantly better than the control group.

Several researchers have investigated interactions between instructional objectives, certain learner characteristics, and type of learning. One study in which sixty-eight students were assigned to the experimental group and fifty-three to the control group found no significant difference in reading comprehension between the

---

62 See above, pp. 28-29.
experimental and control groups. Scores in word-attack skills differed significantly at the .01 level of significance for the experimental group. 63

Nimz 64 investigated the effect on the reading ability of fourth and fifth grade students, using instructional objectives for teaching the specific literal reading comprehension skills. Another purpose was to determine attitude changes that might be attributed to the use of instructional objectives. Half of 406 students were taught by objectives, and the other half made up the nonobjectives group. The instructional objectives group showed no greater gain than pupils whose reading was guided by suggestions provided in basal readers. Positive attitude changes were evidenced by both fourth and fifth graders who used instructional objectives for development of literal comprehension skills. Changes in attitude were more favorable for fifth grade boys than for fifth grade girls or for both sexes in the fourth grade.

Tapscott 65 investigated interactions between instructional objectives, student ability, sex, and whites and blacks. The sample consisted of 120 third grade students enrolled in three elementary

63 Jean E. Hoff, "The Effect of IOX Objectives-Based Reading Test Collections upon Fifth-Grade Comprehension and Word-Attack Skills" (unpublished Doctoral dissertation, North Texas State University, 1974).

64 Irene E. Nimz, "An Investigation of Performance Objectives Concerned with Recognition and Recall of Main Idea, Details, and Sequence for Selected Fourth and Fifth Grade Students" (unpublished Doctoral dissertation, Northern Illinois University, 1974).

schools in an urban school district in North Carolina. The findings revealed that all groups did better when taught by the objectives-based procedure. Of the three groups, the below average ability students showed most significant gains. No significant difference was found between gains of boys and girls. Whites performed significantly higher than blacks.

Hawk\(^6^6\) investigated the effect of the use of instructional objectives on achievement of high school students in social studies. Fifteen social studies classes from four suburban high schools were randomly assigned to three groups. A pretest-posttest control group design was used. The control group teachers had no knowledge of the instructional objectives. Pretest and posttest were split-half, Form A and Form B, constructed by the investigator. The instructional time was three weeks. The Kruskal-Wallis one-way analysis of variance and the Mann-Whitney U Tests were used to evaluate the difference between group means. The findings revealed no significant main effect for instructional objectives.

Summary

An in-depth, historical review of objectives in education was presented in the preceding section. It was noted that one of the earliest and most important sets of educational objectives was introduced in 1918. The need for specifying objectives in terms of learner

\(^6^6\)Duane C. Hawk, "The Effects of Behaviorally Stated Objectives on Student Achievement in an Eleventh Grade American Studies Unit on Immigration," Dissertation Abstracts, XXXVII (July, 1976), 248-249-A.
behavior was expressed as early as the 1920's. Many educators are proponents of the use of instructional objectives, but there is a resisting group which has put to question the value of instructional objectives to teaching and learning. Others have turned to research, suggesting that the value of instructional objectives to teaching and learning is really an empirical question.

Of the thirteen studies reviewed which investigated the hypothesis that student achievement will increase if students make use of instructional objectives during the learning situation, the studies conducted by Duell, Merrill, Kaplan, Daniel, Song, Kaplan and Simmons, and Nassif found significance for instructional objectives. The remaining six studies did not find significance for objectives. These mixed findings attest to the difficulty of generalizing across investigations because of the lack of consistent results. This lack of consistency was also manifest in the Duchastel and Merrill review of studies.

The results obtained from research which addressed the training question showed more consistency in findings across investigations. However, only five studies were reviewed. Of the five, the studies conducted by Bassett and Kibler, Cohen and Hillman, Viel, and Duell found significance for instructional objectives. Generalizability of this research is limited by the small number of studies conducted.

Of the four studies which investigated interactions between instructional objectives, certain learner characteristics, and type of learning, the study conducted by Tapscott found instructional objectives to interact with learner characteristics—the below average ability students showed the most significant gains and whites performed
significantly higher than blacks, but no significant difference was found between gains of boys and girls. Studies investigating interactions between objectives and types of learning found no significant differences between these two variables.

From the evidence reported, there appears to be a need to continue to investigate the hypothesis that if students are provided a short training period before using instructional objectives and use them while learning, their achievement will increase. These results have been promising. There is also the need to continue to investigate the effects of instructional objectives as they interact with learner characteristics, but with improvement in research design. For example, the Tapscott study failed to allow for the effects of SES, hence the hypothesis that whites would perform significantly higher than blacks was supported by the data. If SES had been taken into consideration, that finding probably would have disappeared.

LITERATURE RELATING TO CRITERION-REFERENCED MEASUREMENT IN EDUCATION

History and Distinctions

A corollary to the interest and activity which focused on instructional objectives in the 1960's was the intense interest in development of criterion-referenced measures. Many articles were published, many papers were presented at research association meetings, and many school districts and large publishers developed instructional systems which employed criterion-referenced testing during this same period. According to Davis, however, this concept was not new. Teachers have always, with varying degree of success,
measured the level of performance of their pupils on material or processes that have just been taught by means of tests. In 1864, for example, Chadwick wrote that the Reverend George Fisher had prepared a book called the Scale Book which contained the numbers assigned to each degree of proficiency in the various subjects of examination. . . . The numerical values for spelling . . . are made to depend upon the percentage of mistakes in writing from dictation sentences from works selected for the purpose, examples of which are contained in the Scale Book in order to preserve the same standard of difficulty.67

Recognizing the need for additional measures which tell how proficient students are with respect to objectives of instruction, Thorndike wrote the following in 1913:

The detailed nature and the report to the individual of his school marks were not the vices of the old system. Its vice was its relativity and indefiniteness—the fact already described that a given mark did not mean any defined amount of knowledge, or power, or skill—so that it was bound to be used for relative achievement only . . . . Rivalry with one's own past and with a "bogey," or accepted standard, is entirely feasible, once we have absolute scales for speed at which one can run or the height to which one can jump. Such scales are being constructed. The strength of such impersonal rivalry as a motive, while not as great for the two or three who would compete to lead the class under the old system as that system's emphasis on rivalry with others, is far greater for the rest of the group. To be seventeenth instead of eighteenth, or twenty-third instead of twenty-fifth, does not approach in moving force the zeal to beat one's own record, to see one's practice curve rise week by week, and to get up to a new feat.68

A decade later the concept of mastery learning was introduced into educational discussions. Ebel stated:

More than forty years ago Professor H. C. Morrison at the University of Chicago developed and popularized a method of


teaching based on the mastery of "adaptations" of understanding, appreciation or ability. These, unlike skills, seemed to Professor Morrison not to be matters of degree. "... the pupil has either attained it or he has not." To achieve such an adaptation the instructor should organize his materials into units, each focused on a particular adaptation. He should then follow a systematic teaching routine: teach, test, reteach, retest, to the point of actual mastery.69

Ebel stated that the Morrison concept of mastery learning was popular for awhile but interest declined by 1950.

The recognition of the need for criterion-referenced measures dates back to 1864. Half a century later, criterion-referenced measures were being used, as Davis has observed:

By the 1920s, the Winnetka Plan, the Morrison Unit-Mastery Plan, and the Dalton Plan made use of frequent testing to measure pupil performance of specified skills or tasks at a predetermined level, and programmed instruction made use of short diagnostic tests keyed to each step in the instructional process.70

Thirty years later, Flanagan, noting differences between two measures, made the following observation:

The most basic type of descriptive information obtained from tests refers to the individual's knowledge and ability with respect to the content itself. This information tells us directly what the individual did with respect to the questions and problems set by the test. It contrasts with the other type of information in which the individual's score is described by comparison with other scores obtained on the same test.71

---


70 Davis, Criterion-Referenced Tests, p. 4.

A decade later, Ebel noted differences between content standard scores and normative standard scores when he wrote the following:

By the term content standard test scores in this discussion we will mean a number that indicates the percent of a systematic sample from a defined domain of tasks which an individual has performed successfully.

The word content in the term . . . signifies that the score is based directly on the tasks which make up or provide the content of the test. This is in contrast to normative standard scores, which are based on the relative performance of those who have taken the test.72

Although writers had been making distinctions between two types of measures, the term "criterion-referenced" had not made its way into the literature until Glaser wrote in 1963:

The principal difference between these two kinds of information lies in the standard used as a reference. What I shall call criterion-referenced measures depend upon an absolute standard of quality, while what I term norm-referenced measures depend upon a relative standard.73

While Glaser distinguished between criterion-referenced measures and norm-referenced measures, he also called for alternative approaches to the measurement of learning outcomes. As a result, several writers74 gave credit to Glaser for creating the term "criterion-referenced" and for stimulating interest in the topic. Nitko75 also gave credit to the publication of Glaser's article for

---


74 Davis, Criterion-Referenced Tests, p. 3.

revival of Morrison's concept of mastery learning. In 1968, this concept was enlarged upon by Bloom, according to Gronlund. 76

Writers in measurement and instruction continue to make distinctions between criterion-referenced and norm-referenced measures. The latest was presented by Popham:

A norm-referenced test is designed to determine an examinee's performance in relationship to that of a group of individuals who have previously taken the test. Because we must "reference" an examinee's score to the performance of the norm group in order to make much sense out of it, these tests can be characterized as norm-referenced.

A criterion-referenced test, on the other hand, permits us to determine whether or not an examinee can display a clearly defined set of behaviors, such as a well-delimited type of algebraic skill. The "criterion" to which an examinee's score is referenced is the delimited class of skills which the test is designed to measure. 77

Prior to Popham's latest distinction, Smith delineated more specific distinctions by noting points of contrast between them which included such concepts as validity and reliability and such operations as test construction and item analysis. These are listed in Appendix F.

Popham and Husek noted additional points of contrast and expanded on some of the operations and concepts delineated by Smith.

---


These are listed in Appendix G.

These more specific distinctions were presented in the present study because some in the educational community may not be familiar with the differences criterion-referenced testing makes with respect to these concepts and operations. The detailed analyses should not imply that one test is better than the other, but rather that both are useful and should be used for the purposes specified.

Uses of CRT's

It has been suggested that both CRT's and NRT's are useful to the classroom teacher, and each has its specific use. Garvin suggested that each measure should be used according to the importance of the task:

1. Unless at least one of the instructional objectives of a unit envisions a task that must subsequently be performed at a specified level of competence in at least some situation, CRM is irrelevant because there is no criterion.

2. If public safety, economic responsibility, or other ethical considerations demand that certain tasks be performed only by those "qualified" for them by formal instruction, then CRM of the outcomes of such instruction is clearly indicated.

3. In any instructional sequence where the content is inherently cumulative and the rigor progressively greater, CRM should be used to control entry to successive units. However, if there

---

are several different sequences, differing widely in rigor, NRM is more useful in making appropriate placements.

4. There are certain content areas to which criteria do apply but not everyone need meet them. These are the required subjects; everyone must try to learn them—if only as a matter of public policy—but it is almost preordained that some of them will not.81

Within recent years, the emphasis on individualization of instruction has resulted in the development and implementation of "new" instructional programs which require criterion-referenced testing. Hambleton stated:

Since one of the purposes of individualized programs is to maximize the opportunity for all students to learn, it follows that tests used to monitor student progress should be keyed to the instruction. Furthermore, they should provide information that can be used to measure progress along an absolute ability continuum.82

Day reinforced Hambleton's position, stating that criterion-referenced measurement's purpose is to meet the testing and measurement requirements of the new objectives-based and individualized instructional programs.83

Fremer84 added that criterion-referenced measurement is applicable to survey achievement testing, college selection tests,

---


aptitude batteries, and for classification tests in the armed services. He further stated that the most powerful application is in the area of instructional management where the system allows teachers to combine materials from one or more packaged sets of curriculum offerings and permits them to add to this mixture their own ideas and approaches.

Finally, Green, discussing proposals to alleviate racial and ethnic bias in tests of achievement used in the schools, noted that "CRT's are the best available for use in classrooms by teachers and for measuring progress toward short-term goals."85

Setting Performance Standards

Glaser has suggested that a CRT is deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards.

Performance standards are generally specified by defining a class or domain of tasks that should be performed by the individual. Measurements are taken on representative samples of tasks drawn from this domain and such measurements are referenced directly to this domain for each individual. CRT's are specifically constructed to support generalizations about an individual performance relative to a specified domain of tasks.86

However, Millman, synthesizing some of the literature on establishing standards, took the position that this population of items need not actually exist. What is important, though, is that it is described well enough so that a relatively high degree of agreement can be reached about what


kinds of items are not members of the population. In practice, only a reasonably representative sample of items is required. Garvin, continuing in the Glaser tradition, suggested that different levels of performance standards be established for certain tasks because some of them must be performed at a specifiably higher level than others. Suggesting that performance standards can be set too high or too low, Block found that where mastery is set at 80 to 85 percent correct for each test, students are likely to demonstrate both high achievement and maximal affective development at the end of a unit. Using Block's work as a guide, Gronlund set mastery at 80 percent correct for multiple-choice items, and 90 percent for true-false items. The different percentages which are correct take into account the fact that the student can get a certain percentage of the items correct on a true-false test (50%) and a multiple-choice test (25%) by guessing alone. However, Garvin was suggesting that performance standards be set according to the importance of the task rather than according to the type of item.

Millman reviewed several sources and practices for establishing standards of performance on CRT's:


1. By setting the passing score according to the performance of others.

2. By item content. Simply inspect the items and subjectively decide what raw score the typical "senior" ought to make or exceed.

3. Educational consequences. The score should not be set too high so as to affect future learning.

4. Psychological and financial costs. There should be fewer failures when the costs of failing are high.

5. Errors due to guessing and item sampling. The passing score could be raised to take into account the expected contribution attributed to pure guessing. Alternately, each student's score could be adjusted according to the standard correction-for-guessing formula and this adjusted score compared to the standard.\(^91\)

Hambleton noted that a performance standard score is "typically set to permit the teacher to assign students, on the basis of their performance on each subset of items measuring an objective, into one or two mutually exclusive categories, masters and nonmasters,"\(^92\) but, as noted by Hambleton and Novick,\(^93\) classifying a student into one of several mastery states or categories is the primary problem in criterion-referenced measurement.

---


Guidelines for CRT Development

Several writers have developed guidelines or steps for CRT development. Hambleton's are listed as an example:

1. Task analysis. Puts into perspective the purpose of the test and the characteristics of the examinees.

2. Definition of the content domain. Establishing a content domain that permits explicit items to be written from it.

3. Generation of domain-referenced items. Generate a set of items believed to reflect the domain specified by the objectives.

4. Item analysis. Determine the quality (content validity) of the items either by judging each item by content specialists, or applying empirical techniques frequently used in NRT construction. This last suggestion is not recommended.

5. Item selection. The random selection of items from the domain of valid test items that measure objectives.

6. Establishing test reliability and validity. Content validity is assured if procedures described above are followed closely.

Limitations of Criterion-Referenced Measures

Criterion-referenced measures have limitations, as noted by Ebel in 1970. More recent than Ebel's were those noted by Otto:

---


96 Ebel, "Some Limitations of Criterion-Referenced Measurement."
1. Objectives involving hard-to-measure qualities, such as appreciation or attitudes, may be slighted.

2. Objectives involving the retention and transfer of what is learned may become secondary to the one-time demonstration of mastery of stated objectives.

3. Specifying the universe of tasks (determining critical instructional objectives) to be dealt with is of extreme importance. Good tests will do nothing to overcome the problem of bad objectives. But note that the problem here is no different for norm-referenced testing.

4. Determining proficiency standards can be troublesome. Perfect or near-perfect performance should be required if (a) the criterion objective calls for mastery, (b) the skill is important for future learning, (c) items are objective type and guessing is likely. Less demanding performance may be adequate if any of the three conditions do not prevail.97

Limitations of Norm-Referenced Measures

Most educators are probably aware of the limitations of norm-referenced measures. However, some of the most salient were noted by Otto:

1. The very fact that a test is "standardized" in terms of administration and scoring makes it inappropriate for use with certain groups or individuals. The test may be too difficult or too easy; items may be meaningless or placed at inappropriate levels; directions may be incomprehensible.

2. The test maker's quest for brevity, which unfortunately but pragmatically enhances the salability of tests in some circles, may result in unrealistic time limits and a choice between depth and breadth in sampling. Scores of children who work very slowly but accurately are likely to be meaningless; the sampling of behavior is likely to be superficial or constructed.

3. Group administration may work to the disadvantage of certain individuals. The group situation combined with the standardized conditions may invalidate the test in some instances.

4. The format of the test may restrict the type of items used. A machine scorable format, for example, virtually demands some form of multiple-choice items. Certain behaviors are not adequately sampled with multiple-choice items.

5. Tests at upper grade levels assume ability at lower levels. Finally, Popham has suggested that the need for improving criterion-referenced measurement is crucial, as this area of measurement holds tremendous promise for measurement of learning outcomes. What is needed is a well-financed governmentally initiated project to expand its weak technological base.

Studies Related to the Relationship Between CRT's and NRT's

Duchastel and Merrill have previously suggested that criterion-referenced evaluation "may not be amenable to classical statistical techniques. This should, however, be a minimal factor in determining its usefulness." Nevertheless, five studies were reported in this area and are reviewed below.

Briggs, Stoker, and Scanlon proposed a study to compare performance on a test designed to measure a specific behavioral


100 Duchastel and Merrill, "Effects of Behavioral Objectives on Learning," p. 54.

objective with performance on a test which sampled the content of related lectures and text. The latter test was considered norm-referenced because: behavioral objectives were not stated; the test consisted of random samples of the content of books and lectures; the usual result is a normal distribution of scores; grading is usually based on the curve, and scores are reported as percentiles or letter grades showing each student's ranking in the class, not the absolute level of performance. Subjects were forty-two graduate students in a psychology course taught by Briggs at Florida State University. In order to determine the correlation between the objective-referenced test scores and the content-referenced test scores, a Spearman rank correlation coefficient was calculated based on rank-ordered objective-referenced test scores and content-referenced test scores. The rank correlation was found to be .16, indicating little correlation between performance on one test and performance on the other. It was concluded that this finding could be interpreted as evidence favoring the continued use of CRT's.

Griffin 102 developed instructional objectives and a CRT to measure the objectives, to use in a study conducted in an urban Adult Basic Education Demonstration Center with six volunteer groups of adults. The objectives were developed in the areas of reading, vocabulary, and spelling. Student achievement was also measured by administration of a standardized test. The results of both were compared. The major

findings of the study revealed that there was a significant positive relationship between the students' performance and the standardized instrument and the CRT. The Pearson Product-Moment Correlation Coefficient was computed in determining relationships in Griffin's study. It was concluded that a CRT may offer a supplemental or alternative method of evaluating adult basic learning achievement.

In a previously reported study, Tapscott\textsuperscript{103} found the correlation between the CRT and the NRT to be .50, significant at the .01 level of significance. The correlation for the high ability group was .44; for the average group, .41; and for the low group, .49.

Van Valkenburgh\textsuperscript{104} compared the results of a criterion-referenced instrument and the results of a norm-referenced instrument which had been administered to fourth and seventh grade students and which found relatively high and consistent correlations across tests and subgroups. It was concluded that the relationship between the CRT and the NRT was of sufficient strength that approximate grade equivalent scores can be predicted from a CRT to provide information to educational decision makers pertaining to placement, diagnosis, assessment, prediction, and evaluation.

\textsuperscript{103}\textsuperscript{See discussion of Tapscott's study on pp. 41-42 above.}

Cronis designed a study to determine whether three norm-referenced measures, used individually or in any combination, predict the potential for adaptive behavior of trainable retardates as effectively as a locally developed criterion-referenced measure when compared with a criterion measure. One hundred trainable mentally retarded children were randomly selected from thirty-one public school classes for the trainable mentally retarded in Duval County, Florida. The teachers in each of these classrooms administered the instruments used in the study. A stepwise multiple regression procedure was used to evaluate the data. Scores obtained from the various instruments correlated with the criterion variable when a matrix of intercorrelations was computed. The correlation between the criterion measure and the Cain-Levine Social Competency Scale was $r = .82$; between the criterion measure and the Duval Checklist, $r = .80$; between the criterion measure and the Independent Functioning Subscale of the Adaptive Behavior Scale, $r = .92$; between the criterion measure and the Domestic Occupation Subscale of the Adaptive Behavior Scale, $r = .79$. It was hypothesized that the scores on the Duval Checklist derived from the Criterion-Referenced Measure of Adaptive Behavior would predict better scores on the criterion-referenced measure than would the score on the norm-referenced measures. This was not the case. When a single section from two of the norm-referenced measures were added to the third, the $r$ of the norm-referenced measures with the criterion-referenced measure was raised to $r = .92$.

---

Summary

A brief, historical review of the literature relating to criterion-referenced measurement in education was presented in the preceding section. Although intense interest developed in the early 1960's, there was interest in the concept as early as 1864. Between that time and the early 1960's several writers expressed this need and criterion-referenced measures were actually being used as early as the 1920's. A corollary to this concept was the distinctions noted between CRT's and NRT's as presented by several writers. Detailed contrasts between CRT's and NRT's were presented by Smith, and Popham and Husek, and presented in the current study to show the differences criterion-referenced measurement makes with respect to such concepts as reliability and validity and such operations as test construction and item analysis. This was done because many educators are probably not aware of the differences criterion-referenced measurement makes with respect to these concepts and operations. It was noted that the issue is not to use one or the other, but rather that both are useful tools for the classroom teacher.

Uses of CRT's, the setting of performance standards, guidelines for CRT development, and limitations of both CRT's and NRT's were discussed.

Of the five studies reviewed which compared the results of CRT's with the results of NRT's, four found significant correlations between the CRT and the NRT. Generalizability of this research is limited by the small number of studies conducted.
SUMMARY

This chapter focused on a review of literature related to reading comprehension skills; literature related to objectives in education, and a review of previous studies which have investigated the effects of instructional objectives on learning; literature related to criterion-referenced measurement in education, and a review of previous studies which have investigated the relationship between criterion-referenced tests and norm-referenced tests.

The investigator concluded from the review of the literature related to reading comprehension skills that, even though there is little agreement as to what constitutes reading comprehension by reading authorities, there is considerable agreement among the authors who developed reading comprehension models. The definition of reading comprehension used in the present study was selected because it parallels reading comprehension as specified by the models, and it parallels reading comprehension as specified by the PIRAMID instructional objectives. It follows that the CRT's assess reading comprehension at the literal, interpretive, and critical reading levels.

Although there were intensive emphases on the use of instructional objectives in the 1960's, it was noted that this concept was not new. The need for such emphases had been expressed as early as the 1920's. A small group of educators have expressed reservations about the value of instructional objectives to teaching and learning, while others have turned to research for answers.

The investigator concluded from the studies reviewed that a lack of consistent results in investigations limits generalizability.
Of the twenty-two studies reviewed, only twelve found significance for instructional objectives. Those which addressed the training question found significance for instructional objectives more consistently. However, generalizability is restricted because of the small number of studies reported.

The concept of criterion-referenced measurement also received intensive emphases in the early 1960's as measurement specialists sought additional approaches to the measurement of learning outcomes. Detailed contrasts between CRT's and NRT's were presented, and it was noted that CRT's differ from NRT's on several concepts and operations and that both types of measures are useful tools for the classroom teacher. Of the five studies reviewed which compared the results of CRT's and the results of NRT's, four found significant correlations between the CRT and the NRT.

The investigator concluded from the review of related literature that conducting a quasi-experimental study to investigate the effects of the PIRAMID instructional objectives on reading comprehension achievement for students of middle and low socioeconomic status would make a significant contribution to the growing body of knowledge and research findings on objectives-based instruction. It was also of equal importance to this investigator to compare the results of the CRT and the NRT to ascertain the extent to which they agree in assessment of reading comprehension achievement.
CHAPTER III.

METHODOLOGY

INTRODUCTION

The problem was divided into two parts. The first part investigated the question whether students who were taught reading comprehension skills by means of instructional objectives showed a greater gain than students who were taught the same skills by means of suggestions from a basal reader teacher's guide. Those tested were fourth and fifth grade students of middle and low socioeconomic status. The second part was to ascertain the degree of concordance between results of the criterion-referenced measure and results of the norm-referenced measure.

The purpose was twofold: (1) to compare gains in reading comprehension achievement of an experimental group with gains of a control group, and (2) to compare the pretest and posttest group means on the CRT with those of the NRT to determine the extent of relationship between the two measures. The experimental group used the PIRAMID objectives and the control group used suggestions from a basal reader teacher's guide.

Procedures for testing the hypotheses of the study are presented under sections dealing with the following: (1) population and sample; (2) experimental and control group procedures; (3) research design; (4) sources of data; (5) a description of instruments used; (6) hypotheses; (7) statistical analysis of data; and (8) summary.
POPULATION AND SAMPLE

Identification of the Population

The population was identified by locating school districts in which there were schools using the PIRAMID Instructional System to teach reading comprehension and in which there were similar schools using suggestions from basal reader teacher's guides to teach reading comprehension.

The investigator also made an attempt to identify school districts which were located in close proximity to each other. Three school districts were selected.

Hayward Unified School District, Hayward, California lies at the eastern end of the San Mateo Toll Bridge across San Francisco Bay in Alameda County. The city is approximately fifteen miles southeast of Oakland and has a population in excess of 90,000. Pittsburg Unified School District, Pittsburg, California is located about forty miles east of the metropolitan San Francisco Bay Area in Contra Costa County. The city has a population in excess of 20,000. Yuba City Unified School District, Yuba City, California is located about forty miles north of Sacramento in Sutter County. The city has a population in excess of 13,000.

The study proposal was presented initially in the spring of 1976 to the Director of Program Evaluation and Research for the Hayward Unified School District, the Superintendent of the Pittsburg Unified School District, and the Superintendent of the Yuba City Unified School District. The investigator presented an overview of the entire study and received approval to conduct it in the three school districts.
The population from which the sample was selected consisted of all students enrolled in the fourth and fifth grades in the three school districts. This population represented an experimentally accessible population as the target population is much broader in scope. From the population of this total enrollment, the investigator delimited a more specific group to participate in the study. Delimiting criteria included:

1. Selection of a group of subjects who attended schools which were designated as middle SES and in which the PIRAMID Instructional System was used to teach reading comprehension.

2. Selection of a group of subjects who attended schools which were designated as middle SES and in which suggestions from basal reader teacher's guides were used to teach reading comprehension.

3. Selection of a group of subjects who attended schools which were designated as low SES and in which the PIRAMID Instructional System was used to teach reading comprehension.

4. Selection of a group of subjects who attended schools which were designated as low SES and in which suggestions from basal reader teacher's guides were used to teach reading comprehension.

Randomization of subjects to groups was not feasible with the sample involved since grouping and assignment of subjects were completed before the study was initiated.

The subjects who attended schools using the PIRAMID Instructional System to teach reading comprehension were designated as the experimental group; the subjects who attended schools using suggestions from basal reader teacher's guides to teach reading comprehension were designated as the control group.
Socioeconomic Criteria

The SES of subjects was determined by the occupational status of the parent or guardian of each of the students participating in the study. In the spring of each school year, a State-mandated reading test is administered to all students assigned to the second and third grades in California. The occupation of the parent or guardian of each student tested is furnished by each teacher on the back of the student's test booklet. The test booklets are returned to the California State Department of Education where they are scored and the socioeconomic status is determined for each school. The SES index is determined by using specified criteria.¹ These criteria are shown in Appendix H. Note that the range of the SES index is from 1 to 3.

The SES index and test results are returned to each local school district office the succeeding fall semester. The SES value for each school is the average obtained for all second and third grade pupils tested during the spring. For example, an SES index of 2.30 indicates that the school serves a community with a large percentage of people engaged in professional and semiprofessional occupations. Conversely, an SES index of 1.50 indicates that the school serves a community with a large percentage of people engaged in unskilled and semiskilled occupations.

The California State Department of Education selected occupation as a measure of SES because "a survey has indicated that teachers are more likely to know the occupation of the pupil's parents or guardians

than other indicators of socio-economic status."\(^2\)

For purposes of this study, all participating schools with an SES index below 1.80 were classified as low SES. Schools with an SES index between 2.25 and 2.70 were classified as schools of middle SES.

Selection of Sample

After the initial selection of schools using the PIRAMID Instructional System to teach reading comprehension, the investigator ascertained the SES index for each of the schools from the school district central offices. The SES index of schools using suggestions from basal reader teacher's guides was also obtained from the school district central offices. Schools with similar SES indices were matched, and five principals from the Hayward Unified School District, four from the Pittsburg Unified School District, and two from the Yuba City Unified School District were contacted to obtain approval to conduct the study in their schools.

Eight schools volunteered to participate in the study, which comprised two groups of subjects from "naturally assembled collectives ... as similar as availability permits,"\(^3\) to compare results of using two different treatments for teaching reading comprehension skills during the seven-month period. Table 1 shows each school's SES index. Subjects in the experimental group attended schools in all three school

\(^2\)Reading Test: Second and Third Grades, p. 11.

districts; subjects in the control group attended one school in Hayward and two schools in Pittsburg. The Yuba City control group teachers elected not to participate because of the extensive testing required by the study.

| TABLE 1 |
|---------------------------------c-----~-----------------c-----------|
| SES INDEX OF PARTICIPATING SCHOOLS, 1976-77 |

<table>
<thead>
<tr>
<th>School District</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle</td>
<td>Low</td>
</tr>
<tr>
<td>Hayward</td>
<td>1.76</td>
<td>1.67</td>
</tr>
<tr>
<td>Pittsburg</td>
<td>1.79</td>
<td>1.29</td>
</tr>
<tr>
<td>Yuba City</td>
<td>2.46</td>
<td></td>
</tr>
</tbody>
</table>

Selection of Grade Level

The investigator selected the fourth and fifth grades as the levels of students who were to participate in the study because (1) subjects assigned to the fourth grade level constitute the current group on which SES was determined last spring; (2) this index was also applicable to subjects assigned to the fifth grade level since the SES index for each school was used to describe the SES for subjects at all grade levels; and (3) the measuring instruments used in this investigation require subjects to respond by written responses. Such instruments are more appropriate for subjects at intermediate grade levels than at primary grade levels. The PIRAMID skills continuum begins at kindergarten and terminates at the end of the sixth grade. The grade level
organization of most schools included in the present study terminates at the end of grade five.

A total of 670 fourth and fifth grade subjects in eight schools designated as high or low SES participated in the study. They were partitioned in eight combinations according to grade level, reading approach, and SES. Table 2 shows this distribution.

TABLE 2

NUMBER OF SUBJECTS CLASSIFIED BY GRADE, READING APPROACH, AND SOCIOECONOMIC LEVEL

<table>
<thead>
<tr>
<th>Grade</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Middle SES</td>
<td>Low SES</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>149</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>160</td>
</tr>
</tbody>
</table>

The subjects were assigned to thirty-four self-contained classrooms. The experimental group consisted of twenty classrooms; the control group consisted of fourteen classrooms. Table 3 shows this distribution.
### TABLE 3
DISTRIBUTION OF SUBJECTS

<table>
<thead>
<tr>
<th>Grade</th>
<th>No. of Classrooms</th>
<th>No. of Students</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No. of Classrooms</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>185</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>184</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>369</td>
<td>14</td>
</tr>
</tbody>
</table>

**Teacher Characteristics**

Kerlinger has stated that "schools are known to differ in important characteristics: classes differ, school districts differ, neighborhoods differ, teachers differ."⁴ Speaking to the problem of selection of subjects, Kerlinger further stated:

If a fairly large number of classes are selected and assigned at random to experimental and control groups, there is no great problem. But if they are not assigned at random, certain ones may select themselves into the experimental groups, and these classes may have characteristics that predispose them to have higher mean Y scores than the other classes. For example, their teachers may be more alert, more intelligent, more aggressive.⁵

Since randomization of experimental subjects to groups was not feasible in the present study, an attempt was made to select a large number of classes for both the experimental and control groups to help rule out the possibility of the teacher effectiveness variable causing

---


⁵ Ibid., p. 343.
the experimental treatment to be more effective than in the target population.

EXPERIMENTAL AND CONTROL GROUP PROCEDURES

Evidence of teaching procedures is an important aspect of testing of all hypotheses. The teaching procedures used to teach reading comprehension are presented in detail for experimental and control group teachers in the following discussion.

Experimental Group

Teachers of the experimental group used the PIRAMID Instructional System for implementing instructional objectives, criterion-referenced profiles, CRT's, and a prescription model to teach reading comprehension.

The investigator met with each principal and participating teachers prior to September 27, 1976, and explained the purpose of the study and procedures for data collection. Each group was briefed on pretest and posttest procedures on the CRT, and each teacher received test directions, CRT's, and answer sheets for administering the pretest. At this meeting it was decided to meet five additional times during the seven-month period: October, December, February, May, and June. The purpose of these meetings was to give support and encouragement, answer questions, share information and test results, and provide needed materials for the posttest. The teachers were instructed not to tell the subjects that they were involved in a study. This was done to minimize the "Hawthorne Effect," which is an effect produced in subjects by giving them extra attention. This effect could be confounded with the
effect of the experimental variable. It would be difficult to determine the effect of the treatment variable if steps are not taken to minimize the "Hawthorne Effect." The teachers were told to turn in test instructions, CRT's, and answer sheets following the pretest and the posttest.

Instructional methodology. The PIRAMID Instructional System contains fifty-five instructional objectives and fifty-five CRT's in reading comprehension. For purposes of this study, only forty objectives and CRT's were used, as these were considered to be representative of the reading comprehension objectives and CRT's. The number was reduced also to expedite the study.

Experimental group procedures are explained by a six-step process:

1. The first step in implementing the PIRAMID instructional process is to determine the entry point of each student into the program. Teachers identify each student's pretest score from the previous year's achievement test. This score is then cross-referenced to the PIRAMID objectives by using the conversion table in Appendix I, and the student is entered at the appropriate point.

2. The next step is to determine each student's PIRAMID instructional level. A series of ten CRT's is administered beginning with placement on the objective continuum. The correct responses are indicated on the class profile. If one CRT in the series is not mastered, the student begins his instruction with that objective. If all CRT's are mastered, the next series of ten CRT's is administered.

---

6 PIRAMID Classroom Management System (Yuba City, Calif.: The PIRAMID Consortium, 1972).
3. The next step is to record CRT results on class and student criterion-referenced profiles. The student profile acts as a cumulative record of the student's progress in achieving mastery of all the objectives in the reading skills continuum beginning at kindergarten and terminating at the end of the sixth grade. The profile of the student follows him through the grades so that receiving teachers at the next grade level will have an accurate picture of each student's skill needs and will know specifically where to continue instruction. The class profile provides the teacher with an instant picture of which students in her class demonstrate common skill deficiencies in reading. This information forms the basis for small group and individual instruction consistent with students' identified needs.

4. The next step is to use the PIRAMID criterion-referenced teaching prescription for planning instruction to correct identified skill deficits. A PIRAMID criterion-referenced teaching prescription is furnished each teacher for each objective. The prescription provides the teacher with a list of related developmental skills the teacher should teach in order to insure mastery of the objective, together with a list of instructional materials sources available at the school to teach children the developmental skills. In all cases the final decisions regarding materials to be used and the methods to be followed are left to the teacher implementing the prescription.

5. The next step is to teach the developmental skills necessary to attain the objective. The methodology followed by the teacher accommodates pupil needs and learning styles.
6. The final step is to retest each student on the objective following the instructional period. If the student indicates that he has yet to meet the criteria for the attainment of the instructed objective, he is recycled back through the program for additional teaching. To insure that the student does not lose confidence in his ability to master the instructed objective, he is not recycled more than twice. If, after the second instructional period, he does not seem to be progressing satisfactorily, he is instructed on objectives which he has already mastered. A student who demonstrates mastery of the instructed objective on the first or second test receives instruction on the next objective shown on his profile to be a deficit, and continues in the program until all deficits are met within a given instructional block of time.

The process included tutorial, independent study, small group, and some total group activities. It focused on specific tasks for specific children at specific times.

Instructional aides, volunteers, and tutors assisted with instruction on a daily basis. Reading comprehension was taught daily by teachers of the experimental group.

There are differing mastery requirements for each PIRAMID instructional objective. In reading comprehension, the minimum mastery requirement is 60 percent correct on sequence items and 80 percent correct on other comprehension items.

Teachers did not use steps 1-3 for most students because they had been in the program for several years. The student profile gave a picture of each student's skill needs and the teachers knew specifically
where to continue instruction at the beginning of the school year. The CRT pretest results used in the study were shared with each teacher.

Learning materials. Reading materials were teacher-selected with specific tasks predetermined. A variety of materials was used: State basal reading texts, State supplementary reading texts, district- and county-supplied audiovisual materials and equipment, teacher-made materials, and commercially produced manipulative materials. Many of the materials for teaching reading comprehension are listed in the PIRAMID Prescription Resource Book and coded into each instructional objective. Teachers also used games, tapes, records, kits, puzzles, workbooks, worksheets, programmed materials, magazines, newspapers, library books, encyclopedias, almanacs, yearbooks, and content area textbooks.

Money was made available to teachers for additional instructional materials and supplies.

Control Group

Teachers of the control group used suggestions from basal reader teacher's guides for teaching reading comprehension. The investigator met with each principal and participating teachers prior to September 27, 1976, and gave an overview of the entire study. Questions were answered and procedures for data collection were explained. Each teacher received test directions, CRT's, and answer sheets for administering the pretest. The teachers were instructed not to tell the subjects that they were involved in a study. Again, this was done in order to minimize the "Hawthorne Effect." The teachers were told to turn in test instructions, CRT's, and answer sheets following the pretest and the posttest. In
order to equate the time spent by the investigator in the experimental schools, it was decided to meet five additional times during the seven-month period: October, December, February, May, and June. The purpose of these meetings was to answer questions, give support and encouragement, share information and test results, and provide needed materials for the posttest.

**Instructional methodology.** The teaching approach was basically teacher-to-group. The sequential skill development program was based on the basal reader series. Following the use of a readiness test, the teacher's informal tests, last year's teacher's records, or several days of teacher observation, or a combination of these means of evaluation, the subjects were divided into three groups—the high, middle, and low reading groups. Grouping was consistent over a period of time although individuals within a group might move to another group during the course of the school year.

During the daily reading period the teachers worked separately with each reading group. The lessons began with an introduction to new vocabulary and concepts of the unit or story. Following this introduction, students proceeded to guided silent reading based upon purposes for reading evolved by both teacher and pupils. Oral reading was used as a follow-up to the silent reading purposes. Skill-building exercises comprised a third major step. The final step included supplementary activities for enrichment or motivational purposes.

Unit tests embodied in the basal reader series were given during the period in which the lessons of a specific reader were taught as a continual assessment of growth and progress in reading.
A reading achievement test, provided by the publisher, was administered when all the lessons in the reader had been taught. A student's score showed his standing in relation to others who took the test. A low score indicated that the student needed additional instruction related to the low subtest areas. The scores of other students were used to organize their groups for instruction.

Instructional aides, volunteers, and tutors assisted with group instruction on a daily basis.

**Learning materials.** The reading materials were preselected and embodied in basal reader series. Basal readers used were the Bank Street Readers, the Harper and Row Basic Readers, the Harper and Row Design for Reading Series, the Macmillan Reading Program, the Scott-Foresman Open Highways Readers, and the Ginn Basic Readers. A variety of supplementary materials was used, including workbooks, worksheets, teacher-made materials, commercially produced manipulative materials, games, tapes, records, kits, puzzles, district- and county-supplied audiovisual materials and equipment, programmed materials, magazines, newspapers, library books, encyclopedias, almanacs, yearbooks, and content area textbooks.

Money was available to teachers in control schools for additional materials and supplies.

**RESEARCH DESIGN**

The investigator's approach is described as the Nonrandomized Pretest and Posttest Control Group Design. Schematically, this design

---

is represented by:

\[
\begin{array}{c}
\text{Experimental Group} & O_1 & X & O_2 \\
\text{Control Group} & O_1 & O_2 \\
\end{array}
\]

where the \( O_1 \)'s represent pretest administrations of the CRT and the NRT and the \( O_2 \)'s represent posttest administrations of the CRT and the NRT; the \( X \) means the experimental treatment is given between \( O_1 \) and \( O_2 \) for the experimental group but is not given between \( O_1 \) and \( O_2 \) for the control group; and the \( O_1 \)'s are simultaneous events prior to initiation of the study and the \( O_2 \)'s are simultaneous events at the conclusion of the study. The PIRAMID instructional objectives represent the experimental treatment.

This design was selected because, as Kerlinger has stated, the main strength of the well-planned and well-executed before/after experimental-control group design is that when the control group is added the effects of general historical events and maturation variables, such as respondents growing older, more tired, and the like, should be present in both groups.

That is, if something happens to affect the experimental subjects between the pretest and the posttest, this something should also affect the subjects of the control group. Similarly, the effect of testing . . . should be controlled. For if the testing affects the members of the experimental group it should similarly affect the members of the control group.\(^8\)

The effects of testing were also controlled by use of the non-randomized pretest and posttest control group design because testing was manifest equally in the experimental and the control groups.

\(^8\) Kerlinger, Foundations of Behavioral Research, p. 336.
The groups were equated statistically for pre-experimental differences through analysis of covariance procedures utilizing pretest scores as the covariate. For intact groups analysis of covariance procedures are recommended by Kerlinger, especially to be used with the Nonrandomized Pretest and Posttest Control Group Design. Kerlinger stated:

It is frequently necessary to study groups as they are; subjects cannot be matched or assigned at random. . . . Analysis of covariance is a form of analysis of variance that tests the significance of the differences between means of final experimental data by taking into account the correlation between the dependent variable and one or more covariates, and by adjusting initial mean differences in the experimental groups. That is, the analysis of covariance analyzes the differences between experimental groups on Y after taking into account initial differences in the Y measures (i.e., pretest measures) or differences in some pertinent independent variable. The measure used for the control (pretest measures or measuring on a pertinent variable) is called the covariate.9

SOURCES OF DATA

The following instruments were used as pretest and posttest measures of reading comprehension achievement for both the experimental and control groups.

1. Regular classroom teachers gave the PIRAMID Reading Comprehension CRT10 as the pretest to all fourth and fifth grade subjects the week of September 27, 1976, and as a posttest to all fourth and fifth grade subjects the week of May 2, 1977. The reading comprehension raw scores were compared in the data analysis.

9 Kerlinger, Foundations of Behavioral Research, p. 370.

10 PIRAMID Criterion-Referenced Tests in Reading Comprehension (Yuba City, Calif.: The PIRAMID Consortium, 1974).
Answer sheets for the CRT were developed by the Scan-Tron Corporation, Burlingame, California. The answer sheets are contained in Appendix B.

2. Regular classroom teachers administered the Comprehensive Tests of Basic Skills (CTBS), Form S, Level I,\(^{11}\) as the pretest to all fourth grade subjects the week of October 11, 1976, and as a posttest to all fourth grade subjects the week of May 16, 1977. They administered the CTBS, Form S, Level II,\(^{12}\) as the pretest to all fifth grade subjects the week of October 11, 1976, and as a posttest to all fifth grade subjects the week of May 16, 1977. Normally these norm-referenced instruments are administered on a pretest and posttest basis each school year in the Hayward Unified School District and the Yuba City Unified School District. The reading comprehension subtest grade equivalent scores were compared in the analysis of data.

3. Regular classroom teachers administered the Stanford Achievement Test (SAT), Primary Level III, Form A,\(^{13}\) as the pretest to all fourth grade subjects the week of October 11, 1976, and as a posttest to all fourth grade subjects the week of May 16, 1977. They administered the Stanford Achievement Test, Intermediate Level I, Form A,\(^{14}\) as the

---


pretest to all fifth grade subjects the week of October 11, 1976, and as a posttest to all fifth grade subjects the week of May 16, 1977.

Normally these norm-referenced instruments are administered on a pretest and a posttest basis each school year in the Pittsburg Unified School District. The reading comprehension subtest grade equivalent scores were compared in the analysis of data.

INSTRUMENTATION

The PIRAMID CRT

Validity and reliability of the PIRAMID CRT was established by administering pretests and posttests to over six hundred students of PIRAMID schools on each of the objectives that were measurable by a written test. Pretests were given in September 1973, and posttests were given in February 1974. Validity was reported by an analysis of items which included the point biserial correlation, gain ratio, and sensitivity to instruction indices for each item. Reliability was reported by using the select index for each item and the Kuder-Richardson Formula 20 (KR 20) for each objective. A summary of criteria and findings reported for items and objectives on each index are shown in Appendix J.

The Comprehensive Tests of Basic Skills (CTBS)

Validity of the CTBS was established by having classroom teachers and curriculum and testing specialists write test items for the grades for which the tests were designed; the items were tested in two

tryouts, spring and fall, on national samples; and the vocabulary of the test items was checked for difficulty level and appropriateness by a comparison with the Core Vocabulary of the Educational Developmental Laboratories and the Lorge-Thorndike word list.  

Reliability was reported by using Kuder-Richardson Formula 20 for each grade and level. Interlevel correlations for all tests were computed by use of the Pearson Product-Moment Correlation.

Norming group. The norming sample was drawn from both public and nonpublic schools within the fifty states and the District of Columbia. Seven different regions were represented: New England, Midwest, Great Lakes, Plains, Southeast, Southwest, and the West. The sample comprised 130,000 students in grades K-12. The norming sample was selected to represent the national population in terms of size of city and SES, in addition to geographic region. With respect to ethnic composition, the percentage of black, Spanish-speaking, other minority, and nonminority students in the norming sample were, respectively, 16.7 percent, 7.9 percent, .8 percent, and 74.6 percent, compared with the United States Office of Civil Rights breakdown of the public school population in 1970: 14.9 percent black, 5.1 percent Spanish-speaking, .9 percent other minority, and 79.1 percent nonminority. The percentage of pupils from nonpublic schools for the norming sample on the CTBS was 7 percent.

---


18 Ibid., pp. 5-17.
The Stanford Achievement Test (SAT)

Validity of the SAT was established by having several content specialists write test items for the grades for which the tests were designed. They also conducted first tryouts of items in local school systems in order to verify judgments on performance of various item types and the difficulties of specific items. The items were edited by curricular and measurement experts and by persons with various minority backgrounds. Teachers also shared in the editing process. The items were tested in a national item tryout in the fall of 1970. A team of reviewers gave the items a final examination, with emphasis on curriculum appropriateness, item accuracy, universal quality of items, considerations of possible ethnic and racial bias, adequacy of content coverage by grade level, and clarity of wording. Reliability was reported by two types of reliability coefficients, one in terms of split-half estimates based on odd-even scores corrected by the Spearman-Brown Formula, and the second based on Kuder-Richardson Formula 20.19

Norming group. The norming sample was drawn from both public and nonpublic schools within forty-three states comprising 275,000 pupils in grades K-12. The geographic regions were more generally defined than for the CTBS norming group. The norming samples were drawn from four regions: the Southwest, North Central, Northeast, and West. The norming sample was selected to represent the national population in terms of size of city and SES, in addition to geographic region. With respect to ethnic composition, the percentage of black, Spanish-speaking,  

other minority, and nonminority for the 1973 SAT sample were, respectively, 11.6 percent, 4.6 percent, 1 percent, and 82.8 percent, compared with the 1970 census data on the national population: 11.1 percent black, 4.6 percent Spanish-speaking, 1 percent other minority, and 83.3 percent nonminority. The percentage of nonpublic school pupils was 9 percent.  

The norming groups were identified for both norm-referenced tests used in the study because a different norm-referenced test was used in the Pittsburg Unified School District. Yuba City and Hayward Unified School Districts used the same norm-referenced test. It was assumed by this investigator that the norming samples were similar and were drawn from similar populations.

HYPOTHESES

The following hypotheses were developed to guide the investigation of the problem and to accomplish the purposes of the study:

**Hypothesis 1.** The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the PIRAMID CRT.

**Hypothesis 2.** There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the CRT.

**Hypothesis 3.** There is no difference between the low and middle SES groups with respect to reading comprehension achievement.

---

Hypothesis 4. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the reading comprehension subtest of the NRT.

Hypothesis 5. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the NRT.

Hypothesis 6. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the NRT.

Hypothesis 7. There is no correlation between pretest and posttest measures on reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT.

Hypothesis 8. The correlation between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT is equivalent for the experimental group and the control group.

STATISTICAL ANALYSIS OF DATA

Four measures of reading comprehension were obtained from each of the 670 subjects in the study. Two measures, one CRT and one NRT, were obtained prior to the study, and two measures, one CRT and one NRT, were obtained at the conclusion of the study. The data used were the pretest, posttest, and gain score means on the CRT and the NRT. A
2 by 2 by 2 factorial analysis of covariance was used with pretest scores as the covariate. It was then possible to study the eight combinations according to grade level, reading approach, and SES.

To test the hypotheses, three programs of the SPSS\textsuperscript{21} series were executed: Codebook, Breakdown, and Pearson Correlation, as well as the BMD 05V program of the Biomedical Computer Program Library.\textsuperscript{22} These analyses were available through the Computer Center at the University of the Pacific, Stockton, California. The .05 level of significance was used for this investigation.

**SUMMARY**

This chapter focused on a discussion of methodology. Reviewed were the statement of the problem and purpose, a discussion of the population and sample, experimental and control group procedures, the research design, sources of data, instrumentation, hypotheses, and statistical analysis of data.

The study was conducted in three schools in the Hayward Unified School District, Alameda County, California; four schools in the Pittsburg Unified School District, Contra Costa County, California; and one school in the Yuba City Unified School District, Sutter County, California.

The population from which the sample was selected consisted of all students enrolled in the fourth and fifth grades in the three school


districts although the target population is much broader in scope. A sample of fourth and fifth grade students of middle and low SES who attended eight schools in the three school districts was selected on the basis of specific criteria. This sample was assigned to a Nonrandomized Pretest and Posttest Control Group Design. A total of 369 subjects were assigned to the experimental group and a total of 301 subjects were assigned to the control group. They were partitioned in eight combinations according to grade level, reading approach, and SES. A criterion-referenced instrument and a norm-referenced instrument were administered on a pretest and a posttest to all subjects.

The procedure for the experimental and control groups was described in detail, including (1) instructional methodology and (2) learning materials used. Data analysis was available through the Computer Center at the University of the Pacific, Stockton, California.
CHAPTER IV

FINDINGS OF THE STUDY

INTRODUCTION

The major purpose of this investigation was to compare the mean achievement gains of students who were taught reading comprehension skills by means of instructional objectives with the mean achievement gains of students who were taught the same skills by means of suggestions from a basal reader teacher's guide. Students who were taught reading comprehension skills by means of instructional objectives were designated as the experimental group; those who were taught the same skills by means of suggestions from a basal reader teacher's guide were designated as the control group.

A secondary purpose was to compare the pretest and posttest achievement scores obtained from the CRT with those obtained from the NRT to ascertain the extent of relationship between the two measures.

Reading comprehension achievement was measured by the comprehension subtest of the Stanford Achievement Test and the Comprehensive Tests of Basic Skills, and by forty reading comprehension criterion-referenced tests from the PIRAMID Criterion-Referenced Tests. The purpose of this chapter is to report the findings of the investigation related to the stated hypotheses.
SUBJECTS OF THE STUDY

A total of 670 subjects participated in the study during the 1976-77 school year; 340 were enrolled in the fourth grade and 330 in the fifth grade in three California school districts. The three school districts were the Hayward Unified School District, the Pittsburg Unified School District, and the Yuba City Unified School District. A total of 407 subjects of low SES and a total of 263 subjects of middle SES comprised the total sample. Data related to the subjects consist of pretest, posttest, and gain scores for reading comprehension achievement on both the CRT and the NRT.

STATISTICAL PROCEDURES

The 670 subjects were classified by approach, grade, and SES in a 2 by 2 by 2 scheme, and were grouped in the eight cells formed by this design as shown on page 69 in Chapter III. The pretest scores were subtracted from the posttest scores, giving gain scores. Through an analysis of covariance, it was possible to study the group differences as though they were similar in initial achievement level. Pretest achievement scores served as the covariate in the analysis of covariance procedure. Results of the analysis of covariance are given in terms of main effects for approach and its interactions with grade and socioeconomic level.

Hypotheses 1-6 are presented in terms of the relationships of approach, grade, and SES to each of the measures used. The results for Hypotheses 7 and 8 are presented in terms of the relationships between the CRT and the NRT on the pretest and posttest reading achievement...
scores. This was done for the total group, the experimental group, and the control group.

As stated on page 86, Codebook, Breakdown, and Pearson Correlation, of the SPSS\textsuperscript{1} series, and the BMD 05V program of the Biomedical Computer Program Library\textsuperscript{2} were used to test the hypotheses. All analyses were available through the Computer Center at the University of the Pacific, Stockton, California. The .05 level of significance was deemed adequate for this investigation.

FINDINGS FOR THE HYPOTHESES OF THE STUDY

Hypothesis 1. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when measured by the CRT.

The group means, the mean gains, and the marginal means on the pretest and posttest are summarized in Table 4. These data are reported by approach by grade level and approach by SES. Table 5 presents a summary of the analysis of covariance of the differences between the adjusted means. A significance difference (p < .001) between approaches was evidenced.


TABLE 4

CRT MEANS FOR THE SAMPLE CLASSIFIED BY APPROACH BY GRADE LEVEL AND APPROACH BY SES

<table>
<thead>
<tr>
<th>Approach by Grade Level</th>
<th>Approach by SES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading Approach</td>
</tr>
<tr>
<td>Grade Level</td>
<td>Experimental</td>
</tr>
<tr>
<td>4:</td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td>21.75</td>
</tr>
<tr>
<td>Gain</td>
<td>6.85</td>
</tr>
<tr>
<td>5:</td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>18.51</td>
</tr>
<tr>
<td>Gain</td>
<td>6.00</td>
</tr>
<tr>
<td>Approach Means:</td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>16.70</td>
</tr>
<tr>
<td>Posttest</td>
<td>23.13</td>
</tr>
<tr>
<td>Gain</td>
<td>6.43</td>
</tr>
</tbody>
</table>

(N=670)
TABLE 5

SUMMARY TABLE OF THE ANALYSIS OF COVARIANCE OF THE CRT DATA FOR APPROACH, GRADE LEVEL, AND SES

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>1048.460</td>
<td>1</td>
<td>1048.460</td>
<td>24.257*</td>
<td>.001</td>
</tr>
<tr>
<td>Grade Level</td>
<td>193.457</td>
<td>1</td>
<td>193.457</td>
<td>4.476**</td>
<td>.033</td>
</tr>
<tr>
<td>SES</td>
<td>888.997</td>
<td>1</td>
<td>888.997</td>
<td>20.568*</td>
<td>.001</td>
</tr>
<tr>
<td>Approach by Grade Level</td>
<td>203.719</td>
<td>1</td>
<td>203.719</td>
<td>4.713**</td>
<td>.029</td>
</tr>
<tr>
<td>Approach by SES</td>
<td>202.372</td>
<td>1</td>
<td>202.372</td>
<td>4.682**</td>
<td>.029</td>
</tr>
<tr>
<td>Grade Level by SES</td>
<td>10.868</td>
<td>1</td>
<td>10.868</td>
<td>.251</td>
<td>.999</td>
</tr>
<tr>
<td>Approach by Grade Level by SES</td>
<td>1.123</td>
<td>1</td>
<td>1.123</td>
<td>.026</td>
<td>.999</td>
</tr>
<tr>
<td>Error</td>
<td>28570.243</td>
<td>661</td>
<td>43.223</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* .999F (1, 661) = 11.0

** .99F (1, 661) ~ 3.9
An examination of the data in Table 6 shows that the adjusted mean of 25.02 for the experimental group was larger than the adjusted mean of 22.06 for the control group. The PIRAMID objectives-based approach to teaching reading comprehension was more effective. The actual mean was 23.69 as noted in Table 4 on page 91. There was a raw score difference of 1.33 in the adjusted mean of the experimental group and the actual mean, and a raw score difference of -1.63 in the adjusted mean of the control group and the actual mean. Hypothesis 1 was rejected at the .05 level of significance.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Adjusted Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>369</td>
<td>25.02</td>
</tr>
<tr>
<td>Control</td>
<td>301</td>
<td>22.06</td>
</tr>
</tbody>
</table>

Hypothesis 2. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the CRT.

A significant difference (p < .033) between the fourth and fifth grades was evidenced as noted in Table 5 on page 92. The fourth grade showed a greater gain than the fifth grade subjects. Inspection of Table 7 shows that the fourth grade adjusted mean of 24.24 was larger than the fifth grade adjusted mean of 23.12. The actual mean of 23.69
is shown in Table 4 on page 91. There was a raw score difference of .55 in the adjusted mean of the fourth grade group and the actual mean, and a raw score difference of -.57 in the adjusted mean of the fifth grade group and the actual mean. Hypothesis 2 was rejected at the .05 level of significance.

TABLE 7
CRT ADJUSTED MEANS FOR GRADES FOUR AND FIVE

<table>
<thead>
<tr>
<th>Grade</th>
<th>N</th>
<th>Adjusted Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>340</td>
<td>24.24</td>
</tr>
<tr>
<td>5</td>
<td>330</td>
<td>23.12</td>
</tr>
</tbody>
</table>

Hypothesis 3. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the CRT.

A significant main effect ($p < .001$) was produced for SES, as noted in Table 5 on page 92. The middle SES group showed a greater gain than the low SES group. Table 8 shows that the adjusted mean of 25.40 for the middle SES group was larger than the adjusted mean of 22.59 for the low SES group. The actual mean was 23.69, as shown in Table 4 on page 91. There was a raw score difference of 1.71 in the adjusted mean of the middle SES group and the actual mean, and a raw score difference of -1.10 in the adjusted mean of the low SES group and the actual mean. Hypothesis 3 was rejected at the .05 level of significance.
The finding that fourth grade subjects showed a greater gain on reading comprehension achievement than the fifth grade subjects was not independent of approach. A significant interaction ($p < 0.029$) of approach by grade occurred. Table 9 shows that the mean gain score of 6.85 for the fourth grade experimental group was larger, hence the experimental approach was more effective for fourth grade subjects. Although not significant, the experimental approach was more effective than the control approach even at the fifth grade level. The mean gain score of 6.00 for the fifth grade experimental group was larger than the mean gain scores of 5.96 and 2.70 for the fourth and fifth grade control subjects, respectively. In addition, the relative effects of the approach by grade interaction was more consequential for the fifth grade than for the fourth grade because of the greater difference in fifth grade gain scores. These relationships are shown graphically in Figure 1 on page 96.
TABLE 9

CRT GAIN SCORE MEANS CLASSIFIED BY APPROACH AND GRADE LEVEL

<table>
<thead>
<tr>
<th>Grade</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Means</td>
</tr>
<tr>
<td>4</td>
<td>185</td>
<td>6.85</td>
</tr>
<tr>
<td>5</td>
<td>184</td>
<td>6.00</td>
</tr>
</tbody>
</table>

Figure 1. Graphic Representation of the Approach by Grade Interaction on the CRT
The finding that the middle SES subjects showed a greater gain than the low SES subjects on reading comprehension was not independent of approach. A significant interaction (p<.029) of approach by SES was produced. Table 10 shows that the mean gain score of 9.55 for the middle SES experimental group was larger, hence the experimental approach was more effective for the middle SES experimental subjects. Although not significant, the mean gain score of 5.83 for the low SES experimental subjects was larger than the mean gain scores of the control subjects. In addition, the relative effects of the approach by SES interaction was more consequential for the middle SES subjects than for the low SES subjects because of the greater difference in middle SES gain scores. These relationships are shown graphically in Figure 2 on page 98. There were no significant differences for the first order interaction of grade by SES and the second order interaction of approach by grade by SES on this measure.

<table>
<thead>
<tr>
<th>SES</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Means</td>
</tr>
<tr>
<td>Low</td>
<td>309</td>
<td>5.83</td>
</tr>
<tr>
<td>Middle</td>
<td>60</td>
<td>9.55</td>
</tr>
</tbody>
</table>
Hypothesis 4. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the reading comprehension subtest of the NRT.

The group means, the mean gains, and the marginal means on the pretest and posttest are summarized in Table 11 on page 99. These data are reported by approach by grade level and approach by SES. A summary of the analysis of covariance of the differences between the adjusted group means is reported in Table 12 on page 100. A significant difference (p < .001) between approaches was produced on the NRT.
<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Reading Approach</th>
<th>SES Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td>4: Pretest</td>
<td>3.89</td>
<td>3.84</td>
</tr>
<tr>
<td>4: Posttest</td>
<td>5.22</td>
<td>4.92</td>
</tr>
<tr>
<td>4: Gain</td>
<td>1.33</td>
<td>1.07</td>
</tr>
<tr>
<td>5: Pretest</td>
<td>4.35</td>
<td>5.34</td>
</tr>
<tr>
<td>5: Posttest</td>
<td>5.85</td>
<td>6.35</td>
</tr>
<tr>
<td>5: Gain</td>
<td>1.50</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>5.54</td>
<td>5.61</td>
</tr>
<tr>
<td></td>
<td>1.42</td>
<td>1.04</td>
</tr>
</tbody>
</table>

(N=670)
<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach</td>
<td>34.518</td>
<td>1</td>
<td>34.518</td>
<td>16.585*</td>
<td>.001</td>
</tr>
<tr>
<td>Grade Level</td>
<td>4.761</td>
<td>1</td>
<td>4.761</td>
<td>2.287</td>
<td>.127</td>
</tr>
<tr>
<td>SES</td>
<td>20.017</td>
<td>1</td>
<td>20.017</td>
<td>9.617**</td>
<td>.002</td>
</tr>
<tr>
<td>Approach by Grade Level</td>
<td>3.121</td>
<td>1</td>
<td>3.121</td>
<td>1.499</td>
<td>.219</td>
</tr>
<tr>
<td>Approach by SES</td>
<td>.357</td>
<td>1</td>
<td>.357</td>
<td>.172</td>
<td>.999</td>
</tr>
<tr>
<td>Grade Level by SES</td>
<td>1.842</td>
<td>1</td>
<td>1.842</td>
<td>.885</td>
<td>.999</td>
</tr>
<tr>
<td>Approach by Grade Level by SES</td>
<td>.707</td>
<td>1</td>
<td>.707</td>
<td>.340</td>
<td>.999</td>
</tr>
<tr>
<td>Error</td>
<td>1375.742</td>
<td>661</td>
<td>2.081</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $0.99^2 F (1, 661) = 11.0$
** $0.99^2 F (1, 661) = 3.9$
Table 13 reveals that the adjusted mean of 5.81 for the experimental group was larger than the adjusted mean of 5.28 for the control group. The PIRAMID objectives-based approach to teaching reading comprehension was more effective. The actual mean was 5.57, as shown in Table 11 on page 99. There was a grade equivalent score difference of .24 in the adjusted mean of the experimental group and the actual mean, and a grade equivalent score difference of -.29 in the adjusted mean of the control group and the actual mean. Hypothesis 4 was rejected at the .05 level of significance.

**TABLE 13**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Adjusted Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>369</td>
<td>5.81</td>
</tr>
<tr>
<td>Control</td>
<td>301</td>
<td>5.28</td>
</tr>
</tbody>
</table>

Hypothesis 5. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the NRT.

No significant difference was evidenced between the fourth and fifth grades with respect to reading comprehension achievement on this measure. Therefore Hypothesis 5 was accepted.
Hypothesis 6. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the NRT.

A significant main effect (p < .002) was evidenced for SES. Table 14 shows that the adjusted mean of 5.83 for the middle SES group was larger than the adjusted mean of 5.40 for the low SES group. The actual mean was 5.57. There was a grade equivalent score difference of .26 in the adjusted mean of the middle SES group and the actual mean, and a grade equivalent score difference of -.17 in the adjusted mean of the low SES group and the actual mean. Hypothesis 6 was rejected at the .05 level of significance. There were no significant differences when interactions were taken into consideration on this measure.

**TABLE 14**

<table>
<thead>
<tr>
<th>SES</th>
<th>N</th>
<th>Adjusted Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>407</td>
<td>5.40</td>
</tr>
<tr>
<td>Middle</td>
<td>263</td>
<td>5.83</td>
</tr>
</tbody>
</table>

Hypothesis 7. There is no correlation between pretest and posttest measures on reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT.

The results for testing Hypothesis 7 are presented in terms of the relationship of pretest and posttest measurements between the CRT
and the NRT. The correlation coefficients were calculated for the total group to indicate the magnitude of relationship between the measures. These data are summarized in Table 15.

**TABLE 15**

**CORRELATION MATRIX FOR PRETEST AND POSTTEST BETWEEN THE CRT AND THE NRT FOR THE TOTAL GROUP**

<table>
<thead>
<tr>
<th></th>
<th>CRT 2</th>
<th>NRT 1</th>
<th>NRT 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT 1</td>
<td>.76*</td>
<td>.72*</td>
<td>.70*</td>
</tr>
<tr>
<td>CRT 2</td>
<td></td>
<td>.68*</td>
<td>.74*</td>
</tr>
<tr>
<td>NRT 1</td>
<td></td>
<td></td>
<td>.77*</td>
</tr>
</tbody>
</table>

*Significant at the .001 level. (N=670)

The data suggest a high degree of relationship between pretest and posttest on both measures. There was a substantial correlation between the two measures of reading comprehension achievement. Therefore Hypothesis 7 was rejected.

**Hypothesis 8.** The correlation between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT is equivalent for the experimental group and the control group.

The results for testing Hypothesis 8 are presented in terms of the relationship between the CRT and the NRT on pretest measurements and in terms of the relationship between the CRT and the NRT on posttest measurements. The correlation coefficients were calculated for the
experimental group and for the control group. The relationship between the two measures for the experimental group is summarized in Table 16.

TABLE 16
CORRELATION MATRIX FOR PRETEST AND POSTTEST BETWEEN THE CRT AND THE NRT FOR THE EXPERIMENTAL GROUP

<table>
<thead>
<tr>
<th></th>
<th>CRT 2</th>
<th>NRT 1</th>
<th>NRT 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT 1</td>
<td>.78*</td>
<td>.70*</td>
<td>.68*</td>
</tr>
<tr>
<td>CRT 2</td>
<td></td>
<td>.68*</td>
<td>.76*</td>
</tr>
<tr>
<td>NRT 1</td>
<td></td>
<td></td>
<td>.73*</td>
</tr>
</tbody>
</table>

* Significant at the .001 level. (N=369)

A high degree of relationship occurred between pretest and posttest on both measures for the experimental group. The data suggest a substantial correlation on all observations significant at the .001 level of significance. The differences in correlation coefficients for the experimental group and the total group were quite small. The relationship between the two measures for the control group is summarized in Table 17.
A high degree of relationship was evidenced between pretest and posttest on both measures for the control group. The data suggest a substantial correlation on all observations significant at the .001 level of significance. The difference in correlation coefficients for the control group and the total group were quite small.

Fisher's z_r transformation^3 was employed to test the significance of the difference in pretest correlations of .70 for the experimental group and .75 for the control group. This test was also employed to test the significance of the difference in posttest correlations of .76 for the experimental group and .72 for the control group. Fisher's z_r transformation test is represented by:

\[
z = \frac{z_{r1} - z_{r2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}
\]

where \( z \) represents the critical \( z \) value obtained to determine the significance of the difference in correlation between the experimental group and the control group; \( z_{r_1} \) represents the correlation \( z \) score for the experimental group and \( z_{r_2} \) represents the correlation \( z \) score for the control group. The correlation coefficients of .70, .75, .76, and .72 were converted to the corresponding correlation \( z \) scores for the experimental and control groups, respectively, by using the Transformation of \( r \) to \( z_r \) Table found in Ferguson.\(^4\) \( N_1 \) represents the number of subjects in the experimental group and \( N_2 \) represents the number of subjects in the control group. The details of the analysis are shown below:

1. **Pretest Correlation**

\[
\begin{align*}
\text{z} &= \frac{.867 - .973}{\sqrt{\frac{1}{369 - 3} + \frac{1}{301 - 3}}} \\
&= -1.06 \\
&= -1.36 \text{ value for pretest correlation}
\end{align*}
\]

2. **Posttest Correlation**

\[
\begin{align*}
\text{z} &= \frac{.996 - .908}{\sqrt{\frac{1}{369 - 3} + \frac{1}{301 - 3}}} \\
&= .088 \\
&= 1.13 \text{ value for posttest correlation}
\end{align*}
\]

The computed \( z \) value for pretest correlation was -1.36. The computed \( z \) value for posttest correlation was 1.13. A critical \( z \) value

of 1.96 is required for significance at the .05 level of significance. 

No differential relationship between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT was found for the two groups. Therefore Hypothesis 8 was accepted.

**SUMMARY**

This summary draws together the findings of this investigation on each of the eight hypotheses. The findings as assessed by the CRT revealed that the experimental group subjects showed a significantly greater gain than the control group subjects. When the achievement of the fourth and fifth grade groups was compared, the fourth grade group showed a significantly greater gain than the fifth grade group. When the achievement of the low and middle SES groups was compared, the middle SES group showed a significantly greater gain than the low SES group. Significant interactions of approach by grade and approach by SES occurred. The effect of fourth grade and middle SES subjects having significantly higher mean gain scores in reading comprehension achievement than fifth grade and low SES subjects, respectively, was not independent of the experimental approach. The experimental approach was more effective for the fourth grade and the middle SES groups. No significant differences were produced when grade by SES and approach by grade by SES were taken into consideration.

The findings of this investigation as assessed by the NRT revealed that the experimental group subjects showed a significantly greater gain than the control group subjects. This finding was
consistent with the finding on approach as assessed by the CRT. However, when subjects of fourth and fifth grades were compared, no significant difference occurred. When the achievement of subjects of low and middle SES was compared, the middle SES group showed a significantly greater gain than the low SES group. This finding was consistent with the finding on SES as assessed by the CRT. No significant difference was produced when interactions were taken into consideration on this measure.

The data for Hypothesis 7 indicated a substantial correlation for the total group between pretest and posttest measures across all observations. The data for Hypothesis 8 suggested a high degree of correlation between pretest and posttest measurements for the experimental and control groups. Fisher's $z_r$ transformation test of the difference between correlations was employed to test the significance of the differences between the experimental group pretest and posttest correlation coefficients and the control group pretest and posttest correlation coefficients. No significant differences occurred. The correlation between reading achievement as measured by the CRT and reading achievement as measured by the NRT was not found to be different for the experimental and control groups.
CHAPTER V

CONCLUSIONS, ADMINISTRATION IMPLICATIONS,
AND RECOMMENDATIONS

CONCLUSIONS

The following conclusions based on the findings of this study must be considered within the delimitations of the investigation. The conclusions grouped together the findings of Hypotheses 1, 2, and 3, which relate to the relationships of approach, grade level, and SES to the criterion-referenced instrument, and grouped together the findings of Hypotheses 4, 5, and 6, which relate to the relationships of approach, grade level, and SES to the norm-referenced instrument. The conclusions based on the findings of Hypotheses 7 and 8 are discussed separately. Hypothesis 7 relates to the relationship between the CRT and the NRT on pretest and posttest measurements for the total group. Hypothesis 8 relates to the relationship between the CRT and the NRT on pretest and posttest measurements for the experimental and control groups.

Conclusions Related to Hypotheses 1-3

Hypothesis 1. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when measured by the CRT.
Hypothesis 2. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the CRT.

Hypothesis 3. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the CRT.

The findings did not support Hypothesis 1. On the contrary, the mean gain score of the experimental group was significantly greater than the mean gain score of the control group. It was concluded that the experimental approach was more effective in teaching reading comprehension skills.

The hypothesis that no difference would be found between the fourth and fifth grades with respect to gains in reading achievement was not supported by the findings. The mean gain score of the fourth grade was significantly greater than the mean gain score of the fifth grade. It was concluded that reading comprehension achievement of the fourth grade group was superior to reading comprehension achievement of the fifth grade group.

A significant interaction effect occurred with the experimental approach at the fourth grade. The mean gain score of the fourth grade experimental group was significantly greater than the mean gain scores of the fifth grade experimental group and both grade levels of the control group. Even though not significant, the mean gain score of the fifth grade experimental group was greater than the mean gain scores of both grade levels of the control group. This finding led to the conclusion that the experimental approach was more effective for the fourth
grade level, but this approach also influenced the reading comprehension achievement of the fifth grade group.

The findings did not support Hypothesis 3. In this instance, the mean gain score of the middle SES group was significantly greater than the mean gain score of the low SES group. It was concluded that reading comprehension achievement of the middle SES group was superior to reading comprehension achievement of the low SES group.

A significant interaction effect occurred with the experimental approach and middle SES. The mean gain score of the middle SES experimental group was significantly greater than the mean gain scores of the low SES experimental group and both SES levels of the control group. Moreover, even the mean gain score of the low SES experimental group was greater than the mean gain scores of both SES levels of the control group. This finding led to the conclusion that the experimental approach was more effective for the middle SES group, but this approach also influenced the reading comprehension achievement of the low SES group.

Conclusions Related to Hypotheses 4-6

Hypothesis 4. The reading comprehension mean gain score of the experimental group is equivalent to the reading comprehension mean gain score of the control group when assessed by the reading comprehension subtest of the NRT.

Hypothesis 5. There is no difference between the fourth and fifth grades with respect to reading comprehension achievement when measured by the NRT.

Hypothesis 6. There is no difference between the low and middle SES groups with respect to reading comprehension achievement when measured by the NRT.
The findings did not support Hypothesis 4. In this instance, the mean gain score of the experimental group was significantly greater than the mean gain score of the control group. It was concluded that the experimental approach was more effective in teaching reading comprehension skills as assessed by the NRT.

The hypothesis that no difference would be found between the fourth and fifth grades with respect to gains in reading achievement was supported by the findings. It was concluded that grade level did not significantly influence reading comprehension achievement as assessed by this measure.

The findings did not support Hypothesis 6. The mean gain score of the middle SES group was significantly greater than the mean gain score of the low SES group. It was concluded that reading comprehension achievement of the middle SES group was superior to reading comprehension achievement of the low SES group.

Conclusions Related to Hypothesis 7

Hypothesis 7. There is no correlation between pretest and posttest measures on reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT.

The findings did not support Hypothesis 7. Strong positive correlations were found between these two measures on pretest and posttest assessments of reading comprehension achievement for the total group. It was concluded that the CRT was highly comparable to the NRT as a measuring instrument of reading achievement.
Conclusions Related to Hypothesis 8

Hypothesis 8. The correlation between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT is equivalent for the experimental group and the control group.

The findings supported Hypothesis 8. No difference in correlation was found between reading comprehension achievement as measured by the CRT and reading comprehension achievement as measured by the NRT for the experimental and control groups. It was concluded that the CRT and the NRT were comparable measures of reading achievement.

A summary of the hypotheses, the findings of the study, and the conclusions drawn from the findings is presented in Chart 2 on pages 114 and 115. Reading achievement was significantly enhanced when students were taught reading comprehension skills by the PIRAMID instructional objectives, as assessed by the CRT. This finding was replicated by the NRT. These findings support the view that instructional objectives facilitate learning.\(^1\) Irrespective of teaching methodology, fourth grade had a significant effect on reading comprehension achievement. This finding was not replicated by the NRT. Middle SES had a significant effect on reading comprehension achievement as demonstrated by both measures. The effect of fourth grade and middle SES having a significant effect on reading comprehension achievement was not independent of the experimental approach, which was more effective even for fifth grade and

<table>
<thead>
<tr>
<th>As Assessed by the CRT</th>
<th>Findings</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$ The mean gain score of the experimental group = the mean gain score of the control group.</td>
<td>A significant difference ($P &lt; .001$) occurred in favor of the experimental approach.*</td>
<td>The experimental approach was more effective in teaching reading comprehension skills.</td>
</tr>
<tr>
<td>$H_2$ The mean gain score of the fourth grade = the mean gain score of the fifth grade.</td>
<td>A significant difference ($p &lt; .033$) occurred in favor of the fourth grade.*</td>
<td>Fourth grade reading achievement was superior to fifth grade reading achievement.</td>
</tr>
<tr>
<td>$H_3$ The mean gain score of middle SES = the mean gain score of low SES.</td>
<td>A significant interaction effect ($p &lt; .029$) occurred with the experimental approach at fourth grade.*</td>
<td>The experimental approach was more effective for the fourth grade.</td>
</tr>
</tbody>
</table>

A significant difference ($p < .001$) was produced in favor of the middle SES.*

A significant interaction effect ($p < .002$) was evidenced with the experimental approach at middle SES.

* Hypotheses rejected at the .05 level of significance.
<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Findings</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H4</strong> The mean gain score of the experimental group = the mean gain score of the control group.</td>
<td>A significant difference ($p &lt; .001$) occurred in favor of the experimental approach.*</td>
<td>The experimental approach was more effective in teaching reading comprehension skills.</td>
</tr>
<tr>
<td><strong>H5</strong> The mean gain score of the fourth grade = the mean gain score of the fifth grade.</td>
<td>No significant difference ($p &lt; .05$) was evidenced between the fourth and fifth grades.**</td>
<td>Grade level did not significantly influence reading achievement.</td>
</tr>
<tr>
<td><strong>H6</strong> The mean gain score of middle SES = the mean gain score of low SES.</td>
<td>A significant difference ($p &lt; .002$) was produced in favor of middle SES.*</td>
<td>Reading achievement of the middle SES group was superior to reading achievement of the low SES group.</td>
</tr>
<tr>
<td><strong>H7</strong> There is no correlation between CRT and NRT measurements of reading achievement for the total group.</td>
<td>Significantly high correlations ($p &lt; .001$) occurred between CRT and NRT pretests ($r = .72$) and between CRT and NRT posttests ($r = .74$) for the total group.*</td>
<td>The CRT was highly comparable to the NRT as a measuring instrument of reading achievement.</td>
</tr>
<tr>
<td><strong>H8</strong> The correlation between the CRT and the NRT is equivalent for the experimental and the control groups.</td>
<td>No significant difference ($p &lt; .05$) in correlation occurred between the experimental and control groups.**</td>
<td>The CRT and the NRT were comparable measures of reading achievement.</td>
</tr>
</tbody>
</table>

* Hypothesis rejected at the .05 level of significance.  
** Hypothesis accepted at the .05 level of significance.
low SES groups. These significant and nonsignificant interaction effects were demonstrated by the CRT. There were no significant differences when interactions were taken into consideration as assessed by the NRT. The general conclusion was that the PIRAMID approach was more effective than the basal reader approach in teaching reading comprehension skills to fourth and fifth grade students of middle and low socioeconomic levels. The CRT and the NRT were comparable measures of reading achievement across all measurements.

The sample of 670 students who participated in this study were considered to be representative of the target population. The target population consists of all present and future students, in all grades, in all school districts using the PIRAMID objectives-based approach to teach reading comprehension skills. The inference might be made that the PIRAMID approach may be more effective than the basal reader approach for increasing reading comprehension achievement in the target population. For the 670 students who participated in this research, their grade level and their socioeconomic level influenced reading comprehension achievement. The inference might be made that grade level and SES may influence reading comprehension achievement in the target population. In this instance, the fourth grade and middle SES groups may influence reading comprehension achievement. Moreover, it would be expected that the PIRAMID approach may be more effective for fourth grade students and students of middle SES, although the PIRAMID approach was also more effective than the basal reader approach for the fifth grade and low SES groups. On the basis of this additional evidence, the inference might be made that the PIRAMID approach may be more effective
for all fourth and fifth grade students and students of both socio-economic levels in the target population. The finding that the differential effectiveness of the two approaches of teaching reading comprehension skills was the same for students in grades four and five, and for students of middle and low SES as assessed by the NRT suggests that the PIRAMID approach may be more effective than the basal reader approach for different grade levels and different SES groups in the target population.

**ADMINISTRATIVE IMPLICATIONS**

The findings of this study suggest that schools in the target population might best try merely to teach reading comprehension skills by means of instructional objectives. The present research clearly demonstrated the effectiveness of the PIRAMID objectives-based approach in teaching those skills. The findings and conclusions of this research on teaching approaches suggest important implications for the elementary school principal and faculty. These relate to the kinds of staff development programs that are needed in the future, the faculty's consideration of school organizational patterns which facilitate the use of a diagnostic/prescriptive approach inherent in the PIRAMID instructional methodology, the use of instructional objectives to provide balance in the total curriculum, the establishment of a resource center, the need for parent education, budgetary provisions, and evaluation.
Staff Development Programs

Chall stated that the basal reader approach is "used almost universally by American classroom teachers."\(^2\) To move a faculty from the basal reading teaching methodology to the PIRAMID teaching methodology may not be an easy task for the administrator. The principal will need to be skilled in group dynamics techniques and processes which facilitate change. Future staff development programs should place emphasis on the use of effective teaching methodologies based on empirical knowledge. The faculty will need inservice training on diagnostic/prescriptive teaching and on management of the Instructional System if the PIRAMID approach is adopted.

School Organizational Patterns

The principal and faculty may want to consider other grade organizational patterns if the PIRAMID approach is adopted. Nongraded, multigraded, and continuous progress systems of organizing the elementary school are possibilities. In addition, the principal and faculty may want to consider different methods of grouping students for instruction and different methods of reporting student progress.

Use of Instructional Objectives To Provide Balance in the Total Curriculum

The use of instructional objectives has implications for teaching other basic skill and content subjects. Instructional objectives may be designed for independent study, discovery, and problem-centered

activities, and for other modern instructional strategies utilized in classrooms, media centers, and learning labs in the elementary school.

Bloom's Taxonomy of Educational Objectives\(^3\) will prove useful in evaluating instructional objectives according to the six cognitive areas of knowledge, comprehension, application, analysis, synthesis, and evaluation. Learning will increase when instructional objectives are spelled out in "operational terms."\(^4\)

Establishment of a Resource Center

The principal and faculty will need to plan and develop a resource center which will serve as a central data bank for both published and teacher-devised materials if the PIRAMID approach is adopted. Additional personnel will be needed to staff the resource center, such as a resource center specialist and paraprofessionals.

The Need for Parent Education

The principal and faculty will need to develop and administer effective parent education programs designed to facilitate an understanding of diagnostic/prescriptive teaching and the operationally stated instructional objectives. Communication channels such as Meet-the-Teacher Night, the P.T.A. or other parent/teacher organizations, Open House, and the school newsletter may be used.


Budgetary Provisions

Provisions must be made in the local school budget for purchase of the PIRAMID Instructional System and its support materials and equipment. The principal will need to provide time and funds if instructional materials will be developed by the faculty. In addition, provisions must be made in the district budget for salaries and wages of additional personnel.

Evaluation

The adoption of the PIRAMID approach has implications to the principal and faculty for planning a comprehensive evaluation program and for designing comprehensive evaluation studies. Stufflebeam's Context, Input, Process, Product (CIPP) Evaluation Model\(^5\) may prove useful to the principal and faculty for this process.

RECOMMENDATIONS

The investigator was encouraged by the findings of this investigation. However, since the present study constitutes the only empirical evidence of the effectiveness of teaching by means of the PIRAMID instructional objectives, and since the findings of contemporary research on the effects of objectives-based instruction on student achievement is limited, several recommendations are made for further study:

---

1. To replicate this study to include a larger sample size for the experimental group and the control group and a broader range of SES groups. This was recommended because only two SES groups were used in the present study.

2. To conduct a study similar to the current study in which the amount of investigator supervision would be reduced. The investigator provided equal supervision to both the experimental and control group teachers.

3. To conduct a study on the PIRAMID Math Instructional System to ascertain the effects on achievement of teaching math concepts by means of this approach.

4. To extend this study to two calendar years as a follow-up on the present sample to determine how lasting would be the effects of the experimental approach.

5. The PIRAMID Instructional System requires an extensive amount of testing. This problem came into view frequently during meetings with participating teachers and principals during the 1976-77 school year. It is recommended that the PIRAMID Consortium seek ways to reduce the amount of testing. It might be that more extensive inservice education in this vital area would result in improvement.

6. To establish grade equivalent norms on the PIRAMID CRT's to provide a more reliable basis for comparison of the CRT results with the NRT results.

7. Finally, it is recommended that the California State Department of Education refine its method of determining socioeconomic level of students by including other indicators of SES rather than relying solely on parents' occupation.
SUMMARY

This chapter focused on a discussion of the conclusions drawn from the findings of the study, implications for the administrator, and recommendations for further study. It was concluded that the PIRAMID objectives-based approach was more effective than the basal reader approach in teaching reading comprehension skills; that reading comprehension achievement of the fourth grade and middle SES groups was superior to that of the fifth grade and low SES groups; and that the PIRAMID approach was more effective for fourth grade and middle SES subjects. Moreover, the data suggest that the PIRAMID objectives-based approach was more effective than the basal reader approach for all groups involved in this study. These findings and conclusions readily generalize to the target population. The differential effectiveness of the two approaches was the same for both grades and both SES groups as assessed by the NRT. This suggests that the PIRAMID approach may be more effective than the basal reader approach for different grade levels and different SES levels in the target population. Implications were made for administrators relating to staff development programs, organizational patterns which facilitate use of the PIRAMID approach, use of instructional objectives to provide balance in the total curriculum, establishment of a resource center, parent education, budgetary provisions, and evaluation. Recommendations were made for further study.
Books


Smith, Nila B. American Reading Instruction. Newark, Del.: International Reading Association, 1965.


Wilson, Robert M. Diagnostic and Remedial Reading for Classroom and Clinic. Columbus, Ohio: Charles E. Merrill, 1972.


Articles in Periodicals


Chick, David S. "The Effects on Achievement and Retention of Training Fifth-Grade Children To Use Behavioral Objectives in Self-Instructional Geometry," Dissertation Abstracts, XXXI (December, 1975), 3381-82-A.


Davis, Frederick B. "Research in Comprehension in Reading," Reading Research Quarterly, III (Summer, 1968), 499-545.


Durio, Helen F. "Behavioral Objectives: Where Have They Taken Us?" The Clearinghouse, XLIX (January, 1976), 201-2.


Gordon, Robert A. "Effects of Instructional Objectives on Reading Acquisition and Retention Levels of Students with Varying Mental Ability," Dissertation Abstracts, XXXVI (April, 1976), 6522-23-A.


Hawk, Duane C. "The Effects of Behaviorally Stated Objectives on Student Achievement in an Eleventh Grade American Studies Unit on Immigration," Dissertation Abstracts, XXXVII (July, 1976), 248-49-A.


Petty, Bruce A. "An Investigation of the Effects of Written Behavioral Objectives upon Performance and Attitudes of Students in an Instructional Technology Course at Kansas State University," Dissertation Abstracts, XXXVI (July, 1975), 149-50-A.


**Microfilms**


Green, Donald R. Racial and Ethnic Bias in Achievement Tests and What to Do about It. n.p., March 1974. (ERIC ED 084 285)


Unpublished Materials


Gibson, Terry L. "Effect upon Learning of Student Knowledge and Acceptance of Behavioral Objectives." Paper presented at the Association for Educational Communications and Technology Annual Convention, Atlantic City, 1974.


Publications of the Government and Other Organizations


Comprehensive Tests of Basic Skills, Complete Battery, Level II, Form S,


APPENDICES
APPENDIX A

THE PIRAMID READING COMPREHENSION OBJECTIVES

AND THE CRT

2.081 Given a group of sentences, the student will be able to identify which one could be true. (Literal)

2.082 Given a selection to read, the student will be able to identify it as fact or fantasy. (Critical)

2.097 Given a story to read, the student will be able to identify its main idea. (Literal)

2.103 Given a story orally, the student will be able to identify the cause which effected an event in the story. (Critical)

3.109 Given a passage containing a specific mood and/or feeling, the student will be able to identify the feeling conveyed. (Interpretive)

3.113 Given a paragraph to read, the student will be able to identify the main idea. (Literal)

3.115 Given a short story to read, the student will be able to identify the part which answers the questions of who, what, where, or when. (Literal)

3.117 Given a short story to read, the student will be able to identify the correct title from a group of possible titles. (Literal)

3.118 Given a story and a set of sentences, the student will be able to select the sentence most accurately describing the events of the story. (Literal)

3.119 Given a story to read, the student will be able to place the events of the story in sequence. (Literal)

3.123 Given an oral story, the student will be able to select from a list the best generalization or conclusion. (Interpretive)

3.125 Given a short story, the student will be able to discriminate between a fact and the author's opinion. (Critical)

4.138 Given a list of three statements, the student will be able to select the one which most closely describes the main idea of a given paragraph. (Literal)
4.139 Given a selection, the student will be able to identify a specific fact contained in the selection. (Literal)

4.149 Given a paragraph describing a character in a particular situation, the student will be able to identify the emotion experienced by that character. (Interpretive)

4.159 Given a sentence, the student will be able to identify whether it describes past time or present time. (Interpretive)

5.161 Given a selection to read silently, the student will be able to answer a specific question on its content. (Interpretive)

5.163 Given several headings and a group of items, the student will be able to classify them according to the categorical headings. (Interpretive)

5.165 The student will be able to recall details from a selection read. (Interpretive)

5.171 Given a selection in which only facts are presented, the student will be able to identify a conclusion which may be inferred from the material. (Interpretive)

5.172 Given an exaggerated narrative, such as a tall tale, the student will be able to identify an example of exaggeration. (Critical)

5.174 Given a selection, the student will be able to compare the feelings and attitudes of the main characters. (Interpretive)

5.176 Given a story, the student will be able to identify the author's purpose. (Interpretive)

5.178 Given a selection to read, the student will be able to perceive size, space, or time relationships by answering a set of questions. (Interpretive)

5.180 Given a selection of cause and effect relationships, the student will be able to match each cause statement with its corresponding effect statement. (Interpretive)

5.181 Given a statement, the student will be able to classify it as fact or opinion. (Interpretive)

5.183 After reading a selection, the student will be able to choose a general statement about the selection from a list containing both general and specific statements. (Interpretive)

5.184 Given two story selections, the student will be able to compare and contrast a main character of one story with a main character of the other. (Interpretive)

5.185 Given a reading passage, the student will be able to identify a conclusion drawn from it. (Interpretive)
6.194 The student will be able to read facts and answer questions about the similarities or differences of the things described by the facts. (Interpretive)

6.196 The student will be able to read and interpret facts from a map. (Interpretive)

6.197 The student will be able to read and interpret facts from a chart. (Interpretive)

6.199 Given a paragraph, the student will be able to identify two events or two statements which are inconsistent. (Critical)

6.200 Given a list of words, the student will be able to identify the two that are synonyms. (Interpretive)

6.206 The student will be able to read an article, extract facts, and use these facts in completing a simple outline. (Interpretive)

6.208 Given a facsimile of an article from a newspaper, the student will be able to identify it as an editorial or a news story. (Literal)

6.212 Given a selection to read, the student will be able to choose the best statement of cause and effect from a list supplied. (Critical)

6.220 Given a selection, the student will be able to identify it as: biography, autobiography, fairy tale, myth, or tall tale. (Critical)

5.166 Given a reading selection and a list of events relating to its content, the student will be able to place these events in proper sequence. (Literal)

5.169 Given a list of scrambled chronological events, the student will be able to arrange them in sequential order. (Literal)