Pilot study designed to develop a criterion-referenced test for first grade word-attack skills

Marjorie Elizabeth Vedder
A PILOT STUDY DESIGNED TO DEVELOP
A CRITERION-REFERENCED TEST FOR
FIRST GRADE WORD-ATTACK SKILLS

by
Marjorie Elizabeth Vedder

A RESEARCH PAPER
SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN EDUCATION (READING SPECIALIST)
AT THE CARDINAL STRITCH COLLEGE

Milwaukee, Wisconsin
1973
This research paper has been approved for the Graduate Committee at the Cardinal Stritch College by

(Advisor)

Date [Handwritten: May 1918]
ACKNOWLEDGMENTS

The writer wishes to express her indebtedness to Sister Marie Colette, O.S.F., for her constructive criticism and direction which proved invaluable in the planning and completion of this paper.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Introduction</th>
<th>Scope and Limitations</th>
<th>Summary</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. THE PROBLEM</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>II. REVIEW OF LITERATURE</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>III. PROCEDURE</td>
<td></td>
<td></td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>IV. INTERPRETATION OF DATA</td>
<td></td>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>V. SUMMARY AND CONCLUSIONS</td>
<td></td>
<td></td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
<td></td>
<td></td>
<td>47</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Recognition of Initial Consonant Sounds, (Objective 1)</td>
<td>31</td>
</tr>
<tr>
<td>2. Recognition of Final Consonant Sounds, (Objective 2)</td>
<td>31</td>
</tr>
<tr>
<td>3. Recognition of Medial Consonant Sounds, (Objective 3)</td>
<td>32</td>
</tr>
<tr>
<td>4. Auditory Identification of Initial Digraphs, (Objective 4)</td>
<td>33</td>
</tr>
<tr>
<td>5. Recognition of Initial Digraphs, (Objective 5)</td>
<td>33</td>
</tr>
<tr>
<td>6. Recognition of Initial Blends, (Objective 6)</td>
<td>34</td>
</tr>
<tr>
<td>7. Recognition of Alternate Forms of a Verb, such as, ed, s, and ing, (Objective 7)</td>
<td>35</td>
</tr>
<tr>
<td>8. Recognition of Possessive Forms of a Noun, (Objective 8)</td>
<td>36</td>
</tr>
<tr>
<td>9. Recognition of Plural Forms of a Noun, (Objective 9)</td>
<td>37</td>
</tr>
<tr>
<td>10. Recognition of Compound Words, (Objective 10)</td>
<td>38</td>
</tr>
<tr>
<td>11. Recognition of Phonetic Parts of Words, ew, oo, ow, (Objective 11)</td>
<td>38</td>
</tr>
<tr>
<td>12. Recognition of Long and Short Vowels, (Objective 12)</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER I

THE PROBLEM

Introduction

Criterion-referenced testing has entered the educational field as one of its newest innovations. Research concerning this subject began to appear around 1969, and each succeeding year has seen the number of articles increase. But, as new and innovative as this method may seem, criterion-referenced tests were first introduced to the educational scene about ten years ago by Glaser. In an article in the American Psychologist, Glaser is concerned about measuring individual student achievement. More specifically he states that:

Achievement measurement can be defined as the assessment of terminal or criterion behavior; this involves the determination of the characteristics of student performance with respect to specified standards.\(^1\)

Glaser is concerned with individual performance—with what the student can and cannot do. He states that:

Criterion-referenced measures indicate the content of the behavioral repertory, and the correspondence between what an individual does and the underlying continuum of achievement.\(^2\)


\(^2\)Ibid., p. 520
Glaser has here outlined the framework for criterion-referenced tests. The tests are based on specific behavioral objectives, and the student is measured on his own abilities with reference to himself and not others.

In the light of Glaser's definition, and the increased emphasis on individually-prescribed instruction and the utilization of behavioral objectives in teaching, criterion-referenced testing seems to provide a good method for individual assessment and course evaluation. The child is evaluated as he progresses through the various educational stages, rather than at the end of a unit of instruction. The tests will identify the needs of the individual student and instruction can be planned utilizing the method of ongoing diagnosis and remediation. For as Cronbach has aptly stated:

Evaluation, used to improve the course while it is still fluid, contributes more to improvement of education than evaluation used to appraise a product already placed on the market.¹

The writer of this paper, a first grade teacher, is employed by the Kettle Moraine Area Schools. As a fairly new system, growing and meeting the demands of society and education, increased emphasis has been placed on individually-prescribed instruction and the use of specific behavioral objectives in teaching.

This fact has led the writer to look at criterion-referenced measurement as a tool for evaluating individual performance in reading at the first grade level. No such tests are presently used in the system.

Statement of the Problem

Primary Purpose

The purpose of this paper was to research current publications on criterion-referenced tests and, as a pilot study, design such a test for first grade reading word-attack skills.

Definition of Terms

The following explanations are given in order that certain terms used in the study are clarified.

Criterion-Referenced Test.--A criterion-referenced test is one that is deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards.\(^1\)

Performance Standards.--According to Glaser these are specified by defining a class or domain of tasks that should be performed by the individual student.\(^2\)

Norm-Referenced Test.--A norm referenced test has been described as one that is used to identify an individ-


\(^2\)Ibid., p. 41.
ual's performance in relation to the performance of others on the same measure.  

Validity.--The validity of criterion-referenced measures is determined by the adequacy with which they represent the criterion being tested.  

Reliability.--Popham and Husek state that the typical indices of internal consistency with which one would normally judge reliability in a standardized test are not appropriate for criterion-referenced tests. It is not clear what should replace these indices of internal consistency.  

Item Construction.--The items in a criterion-referenced test are to be constructed in such a way that they are an accurate reflection of the specific criterion behavior.  

Item Analysis.--Item analysis has been used with norm-referenced tests to identify those items that were not discriminating among individuals. In criterion-referenced measurement, an item which does not discriminate need not be eliminated.  


2 Ibid., p. 6.  

3 Ibid., p. 5  

4 Ibid., p. 4  

5 Ibid., p. 6
Behavioral Objective.--A behavioral objective states what a pupil will be able to do after the completion of instruction. It is a precise statement of changes to be effected by the learner.

Scope and Limitations

This pilot study was restricted to forty-two primary students from Zion Public School in Pewaukee, Wisconsin. These students were at various stages of reading growth including the pre-primer, primer, first reader, and completion of first reader levels. The criterion-referenced tests were limited to twelve word-attack objectives for first grade reading skills. Each objective had at least four or as many as ten test items, depending on the type of skill being assessed. The test was constructed utilizing four-choice selected-response items. All of the tests were administered in small groups. The various criterion-referenced tests were tried out on these students to discover possible deficiencies in an item, such as misleading wording or items which did not discriminate effectively.

As this writer could find no useful writing practitioner's handbook for criterion-referenced test construction, the properties of the teacher-made criterion tests do not adhere to the various indices of reliability and validity which one normally expects to find in the more traditional norm-referenced measures.
Summary

This pilot study was carried on with forty-two primary students at various levels of reading ability to ascertain which items in the teacher-made criterion-referenced tests were an adequate reflection of specified behavioral objectives. These items could later be used as part of a first grade package of criterion-referenced tests.
CHAPTER II

REVIEW OF LITERATURE

Introduction

Popham has aptly expressed the state of affairs for criterion-referenced measurement when he said:

Amazing progress has been made in psychological measurement since E. L. Thorndike's hopeful declaration of faith: if anything exists, it exists in quantity; and if it exists in quantity it can be measured. An innovation, however, still has its troubles. It is first ignored, then attacked and criticized, then perhaps tentatively accepted, incorporated in the existing system, and gradually taken for granted.¹

Criterion-referenced tests are one of the newest innovations on the educational scene. Many different definitions of this type of test have been written, and its advantages and limitations have been expounded. Different writers have expressed their philosophy concerning how the various indices of adequacy which one normally associates with norm-referenced tests might be applied to criterion-referenced measures. In addition, the methods of test construction, test usage, and availability of printed tests have all been discussed.

Criterion-Referenced Tests

Not many writers can agree on an explicit definition for criterion-referenced tests. The disagreement usually arises over the number of items per objective, whether or not statistical indices are applicable to this type of test, and what criterion of mastery is to be used.

Following are a few of the various available definitions; Jackson's definition of a criterion-referenced measurement applies only to;

...a test designed and constructed in a manner that defines explicit rules linking patterns of test performance to behavioral referents.¹

Millman, on the other hand, believes that;

...the most fruitful concept of a CR test, also called a domain referenced test, is one whose questions are a representative sample from some identifiable and limited domain.²

Nitko, another writer, thinks that the interpretation of criterion concerns the imposition of an acceptable score magnitude as an index of attainment. In his definition, phrases such as, "working to criterion level" and "mastery" are indicated by obtaining a score equivalent to ninety per cent of the items correct. These tests must fulfill four characteristics. These follow:


The classes of behavior that define different achievement levels are specified as clearly as possible before the test is constructed.

Each behavior class is defined by a set of test situations, (i.e. test items or test tasks) in which the behaviors can be displayed in terms of the proper nuances.

Given that the classes of behavior have been specified and that the test situations have been defined, a representative sampling plan is designed and used to select test tasks that will appear on any form of the test.

The obtained score must be capable of expressing objectively and meaningfully the individual's performance characteristics in these classes of behavior. ¹

All of the definitions seem to revert to Glaser's in that they are based on some sort of objectives and are used to provide information about what an individual can and cannot do. It is because the individual is compared with some established criterion, that these measures can be described as criterion-referenced. These tests also involve the concept of mastery rather than rank position. The meaning of the score is not dependent on comparison to others, as is the case in norm-referenced tests.

Norm-referenced Tests

When comparing norm-referenced and criterion-referenced tests, it is noted that there is one very basic difference. A norm-referenced test is used to identify an individual's status with respect to some established standard

of performance. The individual is compared with some normative group, whereas, in criterion-referenced testing, the individual is compared with himself, with what he can and cannot do. Normative tests are usually constructed to fit the capabilities of the average student. Students who have a history of failure may find the tests that they are forced to take generally too sophisticated. These tests are generally employed where a degree of selectivity is required. The various indices of adequacy, such as reliability and validity, are a necessary ingredient. The normative test writer selects those items that permit variant scores, whereas, the criterion writer wants his items to reflect the criterion behavior of the individual.

Some research has been conducted regarding student reception of the two types of tests. One research writer reports that more positive attitudes are associated with the use of criterion-referenced measures than with norm-referenced ones. Pack reported that:

> The practice of using predominantly norm-referenced measures in assessing student classroom performance instead of evaluating students with respect to some set of criteria or performance standards which directly reflect the objectives of instruction may contribute to student dislike of some subject matter. 📌

Another writer, Garvin, appears to endorse the use of criterion-referenced tests to control entry into suc-

---

cessive instructional units for "any instructional sequence where the content is inherently cumulative and the rigor progressively greater." However, this same writer states that norm-referenced measures should be used "if there are several different sequences differing widely in rigor." He thinks that elementary reading is one area where criterion-referenced testing should not be used.

This writer believes that criterion-referenced measures are an excellent method for assessing reading progress, in that one can ascertain what skills have been mastered before progressing to a succeeding unit. But it is also evident to this writer that:

A good framework for reading will always have more delineated skills, attitudes, and behaviors than are reflected in its written tests.

**Limitations and Advantages**

Like any other innovation, criterion-referenced measurement has both advantages and disadvantages. Ebel has been one of the greatest proponents of its limitations. He finds that there are about three basic limitations. He believes that; (1) they do not tell us all we need to

---


2Ibid., p. 63

know about achievement; (2) they are difficult to obtain on any sound basis, requiring a degree of detail in the specification of objectives that are quite unrealistic to expect and impractical to use; and (3) they are necessary for only a small fraction of important educational achievements. Block disagrees with Ebel's specification of these limitations. Block's reply to limitation one is that even though criterion measures do not tell us all we need to know, they "can provide the only relevant information on pupil learning excellence or deficiency." Block also contends that Ebel's second limitation is weak in that:

...the amount of specificity needed to state the objectives has been greatly exaggerated. Objectives are prespecified primarily to guide the instruction, not to facilitate instrument construction.

Moreover, the formulation of objectives is not as taxing in terms of teacher time and effort as Ebel suggests. Block states that it has been found that the objectives, as well as the evaluation instruments, for each chapter in typical algebra, chemistry, and biology textbooks can be constructed in less than ten weeks, by groups of three to


3Ibid., p. 291
four teachers working about two hours a day.\textsuperscript{1} Ebel's final limitation of criterion-referenced measurement overlooks:

\begin{center}
...the great importance of the learning of those few skills for the fullest development of each pupil's talents, interests, understandings, and appreciations.\textsuperscript{2}
\end{center}

Millman and Brazziel have also indicated some limitations of this type of measurement. Millman states that there is a danger that objectives involving hard-to-measure qualities, like appreciation and attitudes, may be slighted, and in addition, the ability to retain and transfer what is learned may be overlooked.\textsuperscript{3}

Brazziel has listed four disadvantages for this type of measure. He thinks that reporting systems will vary and must be interpreted for children moving into new districts; further work must be done on validation procedures; that comparisons of performances of school districts are not readily available; and the materials for teaching toward specified objectives must always be available if tests are to be valid.\textsuperscript{4}

But, these two writers and Jung have also indicated

\begin{footnotes}
\item[2]\textit{Ibid.}, p. 295
\end{footnotes}
some of the various advantages and strengths of criterion-referenced measurement. Jung states that:

> It will be noted that much of the strength of criterion-referenced measurement depends upon the specification of an adequate set of objectives from which items may be sampled.¹

Brazziel believes that criterion measurement permits direct interpretation of progress in terms of specified behavioral objectives, facilitates individualized instruction, and eliminates a situation where half of American children must always be below median. These tests are usually short and summative allowing checks of student progress at regular intervals. They also enable teachers to compile a comprehensive record of a child's development.² Millman, on the other hand, offers two different advantages. He states that they emphasized proficiency, and the use of this type of test fostered improved student attitudes with less competition for grades.³

**Indices of Adequacy**

Validity, reliability, and standard uses of item sampling become a different type of problem for the construction of criterion-referenced tests. Of all the writers surveyed, none provided an adequate method to measure the

---


validity of the tests as we know it from norm-referenced testing. Some writers offered suggestions for analyzing the reliability of criterion-referenced tests, but none could offer a sound procedure. Item analysis was another area where agreement was difficult to obtain.

In the case of validity, Jug, Popham, Hsu, and Cox all think that criterion-referenced measures are validated primarily in terms of the adequacy with which they represent the criterion. These tests must provide information in terms of specific behavior. This type of content validity seems to be the more logical index for criterion-referenced tests than the traditional empirical validity of standardized tests. Popham and Husek believe that validity is irrelevant because the meaning of a score is not dependent upon comparisons with others.¹ Jung states that:

Extensive tryouts of items and standardization are not required, since the test acquires its validity primarily in terms of its relationship to the behaviors delimited by the criterion.²

The question of reliability for criterion-referenced tests also yielded varied viewpoints on the part of the writers. Some had suggestions for determining the reliability of these tests, but again none really agreed.

---


Jackson thinks that:

One way reliability may be analyzed is through comparison of the inferences made for a group of individuals on one form of the test with the inferences yielded by an alternate form developed with identical procedures. An index of agreement between the two forms in classifying the individual tested—perhaps a contingency co-efficient—could be used as an index of "reliability" of the measurement procedures....One further procedure for empirically evaluating certain criterion-referenced tests is a form of a Guttman scale.1

Jackson seems to lean toward a test re-test method of securing reliability.

Edmonston, who has constructed a decision model designed to determine the reliability and validity of criterion-referenced measurements, also looks at the test-retest method as a logical method of obtaining reliability.2 In contract, Hsu does not believe that this method is adequate, since appropriate test items may not be chosen. He states that, "to apply classical reliability formulas for a criterion-referenced test by disregarding different behavioral objectives within a test is evidently undesirable."3

In considering the above authors, it appears that none could provide a workable formula for establishing

reliability in teacher-made tests. Shavelson, Block and Ravitch, appear a bit more realistic when they indicate that a criterion-referenced reliability coefficient may be unnecessary. They do, however, believe that:

One solution to the reliability problem posed by common usage of CR tests is the following: The test should be divided into subscales with a criterion for each subscale. ... Subscale reliability will be lower than total score reliability, but it will provide the information required for diagnosing individuals at the subscale level. ... Persons who clearly reach criterion on a subscale should be classified "mastery." Persons below criterion, say one SEM should be classified "non-mastery" and remediation prescribed. And persons between "non-mastery" and "mastery" should receive additional parallel items to determine their appropriate classification.¹

These last writers seem to offer a more workable plan to use with teacher-made tests.

Item analysis has been used with norm-referenced tests to indicate which items are not discriminating among individuals. In a criterion-referenced test, the items are constructed to reflect the accurate criterion behavior. The test-maker must analyze the items to make sure they are truly representative. Hsu states that a good test item:

Does not only discriminate pre and post-learning, it is also a function of the item to allow the individual to answer correctly if he masters the criterion behavior represented by the item and

answer incorrectly if he actually does not master it; regardless of whether the test is administered before or after formal instruction.¹

When the test writer finds an item which doesn't discriminate, he need not eliminate it. It should remain in the test as long as it is a reflection of an important attribute of the criterion. When a negatively-discriminating item is found, the writer should be suspicious of it and try to detect its flaws through careful analysis.²

In the case of teacher-made criterion-referenced tests, the problem of item analysis on a large scale becomes almost impossible. Popham and Husek suggest that the only way that this could be accomplished was with the aid of data-processing equipment.³

As none of the writers could provide an adequate practitioners' handbook for writing criterion-referenced tests, the test-writer should work to make his items accurately sample the range of criterion behavior that they are testing.

**Test Construction**

The first step in the construction of a criterion test is the selection of a specific set of behavioral


³Ibid., p. 3.
objectives. In the reading program, the use of prespecified instructional objectives, "provides a key to maximally effective classroom instruction when put into operation in the form of criterion measures..."\(^1\) Farr has described the four steps in the establishment of adequate behavioral objectives for this type of test. The objectives must be consistent with the writer's definition of reading; arbitrary criterion levels for specific objectives must not be established merely because a behavioral objective is supposed to have a criterion level; one must not develop subobjectives, and sub-sub-objectives "until the behavior being developed is so badly fractured that it is no longer recognizable," and finally one must not let the instructional procedures become the objectives.\(^2\)

Once the items have been identified, the individual test items must be constructed. The type of test question has also been discussed by various writers. Niedermeyer and Sullivan conducted research to ascertain what type of test item was the most adequate for a criterion-referenced test in reading. Three types of tests were used: (1) individually-administered, constructed-response tests; (2) group-administered, selected-response tests with three

\(^1\)James H. Block, School Review, p. 294.

choices per item; and (3) group-administered, selected response tests with four choices per item. The results of this study indicate that:

...the three-choice, selected-response tests often utilized in programs of this type do not provide an accurate indication of end-of-year achievement for many children, and their continued use is not recommended. The four-choice, selected-response tests used during the year provided an accurate estimate for a substantially larger proportion of the class. The constructed-response tests were well-received by teachers and may be as useful as the four-choice tests, but this cannot be concluded on the basis of the present study.¹

Several writers disagree and believe that multiple-choice items should not be used to test objectives. Johnson and Kress state:

To use the multiple-choice aided recall format characteristic of standardized norm tests is contradictory to the basic purpose to be accomplished—to find out if the testee can complete the reading test successfully.²

Prescott also frowns upon the use of multiple choice items. Prescott states:

...the use of the multiple-choice items (or its variant) is questionable when the criterion-referenced approach is paramount, since it sets up a loaded situation to which the pupil responds.³


Neither writer has, however, supported their statements with research in connection with criterion-referenced tests.

The various writers cannot agree on the number of items that should be used for each objective. Most seemed to favor about five. But, the point was made that unless the list is to become unmanageably long, very few items can be devoted to any one specific behavior. "The number of items measuring each objective in the test and the criterion of mastery are crucial problems."¹ In the pilot study for the Prescriptive Mathematics Inventory, the forty items which represented the range of forty areas in mathematics that were to be assessed were first selected. For twelve of the forty items, twenty alternate forms were written, and for the remaining twenty-eight items, there were five alternate forms. It was found that these tests were too long and that the students became tired and bored.²

Most writers were in agreement that some standard of mastery should be established for each criterion-referenced instrument. Hsu has a rather flexible, sound suggestion when he states:

A mastery level should be determined for each criterion behavior (or objective) rather than


²Ibid., p. 4.
judging a group of objectives as a whole. The mastery level will not necessarily be the same for every objective in a test.

Millman feels that perfect or near-perfect performance of at least ninety per cent should be required if the objective is worded in such a way that near mastery is expected. Mastery should be required if the skills are deemed important for future learning. Niedermeyer and Sullivan looked for a ninety plus per cent accuracy, while Nitko favored a score equivalent to ninety per cent, correct. The NEA Bulletin also states that an acceptable level of performance on criterion-referenced tests should be ninety per cent, and all pupils should be expected to perform at this level, or relearning is necessary.

Uses of Criterion-referenced Tests

Criterion-referenced tests can be constructed and used both by teachers and by professional test-makers. When adequately constructed, these tests reflect the actual steps in the learning process, and when they are appropriately used they can offer real promise for instructional guidance. Criterion-referenced instruments may be used by

1 E.C. Hsu, ERIC Document ED 050-139, p. 8.
2 Jason Millman, Phi Delta Kappan, p. 229.
teachers in parent conferences to communicate to them precisely what their child can do. These tests can also become a basis for a pupil reporting system. Criterion-referenced tests can be used to prescribe individual instruction on a day-to-day basis, and in addition, they may be a tool for accountability. In some states, such as Florida and New Jersey, criterion-referenced tests have been used in their statewide assessment programs.1

Commercially made criterion-referenced tests are not readily available at present. Two of the earliest ones to be published are the Prescriptive Mathematics Inventory (Grades 4-8) issued in 1971, and its companion series the Prescriptive Reading Inventory (Grades 1-5).2 The PMI and PRI present multiple-choice questions and indicate pupil mastery or non-mastery for specific objectives. The PMI covers 345 specific objectives with one question each, while the PRI covers ninety objectives with three or four test questions each. Science Research Associates has issued, thirty-two diagnostic criterion-referenced tests called "Probes."3 These tests represent the core of the

2G.E. Roudabush and D.R. Green, ERIC Document ED 050-144, p. 3.
new mathematics material for *Diagnosis, An Instruction Aid, Level B* for grades four thru six. Random House has also published a criterion reading program which is an individualized learning management system.\(^1\) Harcourt, Brace and Javanovich has completed most of the developmental work on an extensive series of criterion-referenced tests called the *HBJ Reading Assessment System*.\(^2\) This year it is being field tested in grade four classes of three Michigan school districts. In addition, the *Wisconsin Design for Reading Skill Development*\(^3\) and the *Fountain Valley Support System*\(^4\) may be thought of as incorporating the philosophy of criterion-referenced measurement.

**Summary**

Criterion-referenced tests have many and varied definitions depending upon the particular author one is reading. Yet, all seem to refer to Glaser's original definition, in that they are "directly interpretable in terms of specified performance standards," and all are concerned with what an individual can and cannot do. This


\(^2\)Ibid., p. 40.

\(^3\)Wisconsin Research and Development Center for Cognitive Learning.

\(^4\)Richard L. Zweig Associates, Inc.
fact sets them apart from norm-referenced tests, which identify an individual's performance in relation to others. Criterion-referenced tests, like all tests, have advantages as well as disadvantages. Most writers, however, believe that their advantages far outweigh their disadvantages.

The various indices of adequacy is another area which has provoked a good deal of comment. Considering the viewpoints of all of the writers leads one to accept the most workable concepts. The more practical definition held by several writers concerning validity stated that, the validity of criterion-referenced measures is determined by the adequacy with which they represent the criterion being tested. Moreover, the concept of reliability, as a typical indice of internal consistency in a standardized test, is not appropriate for criterion-referenced tests. The items in a criterion-referenced test are to be constructed in such a way that they are an accurate reflection of specific criterion behavior, and items which do not discriminate need not be eliminated.

In teacher-made criterion-referenced tests, a set of behavioral objectives is designated from which the various tests are designed. The four-choice selected-response type of question seems the most reliable. These tests require mastery at the ninety per cent level.

Not many commercial criterion-referenced tests have been published, but more and more are beginning to make
an appearance. The published tests as well as the teacher-made ones, have become a framework for individually prescribed instruction or a tool for accountability.
CHAPTER III

PROCEDURE

Introduction

A primary purpose of this paper was, as a pilot study, to design a criterion-referenced test for first grade reading-word-attack skills. The tool will be used as a measure for evaluating individual performance in reading at the first-grade level. This pilot study was restricted to forty-two primary students from Zion Public School in Pewaukee, Wisconsin. These students were at various stages of reading growth including the pre-primer, primer, first reader and completion of first reader levels.

Procedure

The first step in the construction of the criterion-referenced test was the selection of a specific set of behavioral objectives for the twelve first-grade word-attack skills being measured. The objectives included the following skills: consonant sounds in the initial, final, and medial positions; diagraphs; selected blends; verb endings, such as ed, s, and ing; the possessive form of nouns; the plural form of nouns, compound words, the phonetic parts of words, ew, oo, and ow; and short and long vowels. Once the objectives were written, test items had to be selected for
each objective. Each objective had a minimum of four and a maximum of ten test items. At times, a ten-item test was constructed to combine forms of a skill, as in Objective 7, where alternate forms of a verb, such as ed, s, and ing, were used on one test. It was believed that too many tests might tend to tire the children. The tests were written utilizing four-choice selected-response items.

Administration

The tests were administered by the writer during January and March, 1973, at Zion Public School. On Tuesday, January 8, at ten o'clock, fifteen children at the pre-primer level were given the criterion test for Objective 1. On this same day, immediately following the first test, seventeen children at the primer level took Objective 2. On the next day, January 9, fifteen primer children were given Objective 3. Objective 4 was administered to seventeen primer children on the fifteenth of January at eleven o'clock. Objectives 5 and 6 were taken by fourteen primer children the following morning at ten o'clock. On the sixteenth of March, beginning at ten o'clock, Objectives 7, 8, and 9 were administered to small groups of primer and first reader children. At eleven o'clock on the twentieth of March, Objectives 10, 11, and 12 were taken by ten children who had completed the first reader. Once the tests were given, they were corrected by the writer and data from each objective were compiled in a short table which indicated the
number of correct and incorrect responses for each item.

Summary

The criterion-referenced test was compiled utilizing twelve objectives for first-grade word-attack skills. The tests were administered by the writer to small groups of primary students during January and March of 1973. Results of the tests were then tabulated in the form of an item-analysis.
CHAPTER IV

INTERPRETATION OF DATA

Introduction

The results of the twelve objective criterion-referenced tests used in this pilot study were compiled in short tables. The tables were done in the form of an item analysis indicating the number of correct and incorrect responses per item. Also indicated in the Table were the number of children taking the test and their level of reading achievement. The writer believed that if five or more children missed an item, it might not be a valid one to retain in a subsequent test, or if many children did poorly on an objective, reteaching and retesting would be necessary.

Results

Table 1 indicates the results for Objective 1, which dealt with the ability to recognize consonant sounds in the initial position. Results on Objective 1 are shown in Table I. Fifteen children, the total number of pupils reading at the pre-primer level, took the test. The results show that a majority of children have mastered this skill. Only one child missed items three and four respectively, and item five was failed by two children.
TABLE 1  
RECOGNITION OF INITIAL CONSONANT SOUNDS (OBJECTIVE 1)  
(N=15 at Pre-Primer level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>2</td>
</tr>
</tbody>
</table>

Objective 2 was taken by seventeen children, the total number at the primer level. Results on Objective 2 are shown in Table 2.

TABLE 2  
RECOGNITION OF FINAL CONSONANT SOUNDS (OBJECTIVE 2)  
(N=17 at Primer level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>2</td>
</tr>
</tbody>
</table>

This objective, which tested consonant sounds in the final position, was mastered by most of the children. Item three proved the most troublesome. The children were unable to supply the correct consonant for the word "trip." This
item probably should be changed on subsequent forms of the test.

Fifteen children at the primer level took Objective 3, which dealt with consonant sounds in the medial position. Two children were absent due to illness. Results on Objective 3 are shown in Table 3.

**TABLE 3**  
RECOGNITION OF MEDIAL CONSONANT SOUNDS (OBJECTIVE 3)  
(N=15 at Primer level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

Again the children did fairly well on this objective. Item five gave the most trouble, but this was due to the fact that the children were unable to decipher the drawing of the tulip. This picture should be corrected on a new form of the test.

Objective 4 was taken by seventeen primer children. Table 4 shows the results of this Objective.
TABLE 4
AUDITORY IDENTIFICATION OF INITIAL DIGRAPHS (OBJECTIVE 4)
(N=17 at Primer level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>0</td>
</tr>
</tbody>
</table>

For this criterion test, the children heard a word and they had to choose the correct beginning sound from a list of four digraphs. All children, except one, experienced mastery for this objective.

Fourteen of the primer children took Objective 5, which also dealt with digraphs. Three of the primer children were absent that day. Results on Objective 5 are shown in Table 5.

TABLE 5
RECOGNITION OF INITIAL DIGRAPHS (OBJECTIVE 5)
(N=14 at Primer Level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>
The children had to discover a dissimilar beginning digraph in a set of four words. All but two items were correctly answered on this criterion test.

Fourteen children at the primer level took Objective 6. Again, three children were absent. Results on Objective 6 are shown in Table 6. This test dealt with blends. The children were asked to supply the correct beginning blend for words. The majority of the children did rather poorly on this Objective. None of the children got all of them correct. In analyzing the items, it does not appear that the words are too difficult for the primer level. This writer believes that reteaching should be done on this Objective, and then the test should be readministered. If the test continues to give difficulties, new items should be constructed.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>
Objective 7 was taken by nine primer and fifteen first reader children. These first reader children are the same children who were at the primer level in January. The nine primer children were from the original group of fifteen pre-primer children. Three of the original fifteen had not gone into primers and the other three were absent. Results for this objective are shown in Table 7. This test dealt with alternate forms of a verb, such as ed, s, and ing. Again, the children did poorly on this objective. One child had one wrong and two children missed two items, but the rest of the scores were rather varied. The sentences for this test were taken from various primer and first reader books, so the children should have mastered the basic vocabulary. The poor scores indicate that reteaching and retesting are necessary.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>
Possessive forms of a noun were dealt with in Objective 8. Fifteen first reader children took this test. Results on Objective 8 are shown in Table 8.

### TABLE 8
RECOGNITION OF POSSESSIVE FORMS OF A NOUN (OBJECTIVE 8)
(N=15 at First Reader Level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>4</td>
</tr>
</tbody>
</table>

All did a fairly good job on it. Item four gave the most difficulty, with five children unable to supply the word "Daddy's." This is a rather simple sentence, so this writer suggests keeping it in the test and trying it out on a different set of children.

Objective 9 was taken by fifteen children at the first reader level. Results on Objective 9 are shown in Table 9.
TABLE 9
RECOGNITION OF PLURAL FORMS
OF A NOUN (OBJECTIVE 9)
(N=15 at First Reader level)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

This objective dealt with plural forms of a noun. All the children did poorly on this test. Here again, it would appear that reteaching of the skill is necessary, and then retesting.

Compound words were the concern in Objective 10. Results on Objective 10 are shown in Table 10. Ten children at the completion of the first reader level took the test. The children are not as strong as they could be in this area. Item five gave the most difficulty. The sentence called for the compound word, "yourself," and all four children who got it incorrect, supplied "your way." On subsequent forms of this test a different word should be substituted for "way."
TABLE 10
RECOGNITION OF COMPOUND WORDS
(OBJECTIVE 10) (N=10
Completion of First Reader)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Ten children at the completion of the first reader level took Objective 11. Results on Objective 11 are shown in Table 11.

TABLE 11
RECOGNITION OF PHONETIC PARTS OF WORDS, EW, OW (OBJECTIVE 11)
(N=10 at Completion of First Reader)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>
This test considered the phonetic parts of words. All of the children did well on this objective, with three being the maximum number of items missed on this measure.

The children really understood long and short sounds of vowels, in Objective 12. Table 12 indicates the results for this Objective.

**TABLE 12**
RECOGNITION OF LONG AND SHORT VOWELS (OBJECTIVE 12)
(N=10 at Completion of First Reader)

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Number Correct</th>
<th>Number Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

Ten children who had just completed the first reader level took this test, and all of them experienced good results.

**Summary**

Short tables were compiled to tabulate the results of the twelve criterion-referenced tests. If five or more children missed an item, it was judged to be a poor one. If all the children did poorly on an objective, reteaching and retesting were recommended by the writer.
CHAPTER V

SUMMARY AND CONCLUSIONS

Introduction

Criterion-referenced tests have entered the educational field as one of its newest innovations. After investigating the many and varied definitions for these tests, most educators seem to agree with Glaser's original definition that they are, "directly interpretable in terms of specified performance standards."¹ All criterion-referenced tests are concerned with what an individual pupil can and cannot do. This fact sets them apart from norm-referenced tests, which identify an individual's performance in relation to others. In teacher-made criterion-referenced tests, a set of behavioral objectives is designated from which the various tests are designed.

This paper was developed as a pilot study to design a criterion-referenced test for first-grade word-attack skills. The tool was used for evaluating individual performance on twelve word-attack objectives. The study was restricted to forty-two primary students from Zion Public

School in Pewaukee, Wisconsin. These students were at various stages of reading growth, including the pre-primer, primer, first reader and completion of first reader levels. The tests were administered during January and March of 1973, by the writer, and the results were tabulated. If five or more children missed an item it might not be a valid one to retain on a subsequent test, or if many children did poorly on an objective, reteaching and retesting would be necessary.

Conclusions

Results of the pilot study indicate that the children did well on consonant sounds in the initial, final, and medial positions. But in Objective 3, the drawing of the tulip should be changed. Also, all of the children did a good job on digraphs, compound words, phonetic parts of words, and long and short vowels. In Objective 10, dealing with compounds, a different word should be substituted for "way". However, the children did poorly on blends, alternate forms of a verb such as ed, s, and ing, possessive forms of a noun, and plural forms of a noun. Upon investigation of the test items for these objectives, it appears that reteaching and retesting are necessary before any decision should be made regarding the items. If, after reteaching and retesting, the children still do poorly, then the items should be reconstructed. The results of the study show that a criterion-referenced test is a good instrument for determining an individual's competence in a particular skill.
Summary

Criterion-referenced measurement, as an educational innovation, appears to be an excellent device for diagnosing an individual's progress in a specific area. Results of the pilot study indicate that various inadequacies are revealed by this type of test. For a teacher in this age of accountability and individually-prescribed instruction, criterion-referenced measurement seems to be the answer to a strong need. As long as there are children to teach, there will be children to test, for as Thorndike has stated, "if anything exists, it exists in quantity, and if it exists in quantity, it can be measured."\(^1\)

\(^1\)W. James Popham, *Criterion-Referenced Measurement*, p. v.
BIBLIOGRAPHY

Books


Articles In Journals


Jung, Steven M. "Criterion Measures For Educational Incentives." Psychology In The Schools, VIII (April, 1971), 121-128.

Livingston, Samuel A. "Reply to Shavelson, Block, And Ravitch's 'Criterion-Referenced Testing: Comments On Reliability.'" Journal Of Educational Measurement, IX (Summer, 1972), 139.


Tyler, Ralph W. "Testing For Accountability." Education Digest, XXXVI (March, 1971), 12-14.

Bulletins


ERIC Documents


Test Developers

Wisconsin Research and Development Center for Cognitive Learning, Wisconsin Design For Reading Skill Development, 1404 Regent St., Madison, Wisconsin 53706.

Zweig, Richard L., Fountain Valley Support System, 20800 Beach Blvd., Huntington Beach, California, 92648.
APPENDIX I

Following is a copy of the criterion-referenced test utilizing twelve objectives for first grade word-attack skills. Each criterion test begins with a statement of the specific objective.
Objective 1.

Given specified words the child will recognize consonant sounds in the initial position.

Directions: Circle the word that best completes each question. (This may be done individually or in a group, and the teacher may read each sentence.)

1. What word begins like man?
   1. not  2. well  3. make  4. open

2. What word begins like will?
   1. Our  2. now  3. want  4. saw

3. What word begins like jump?
   1. play  2. going  3. you  4. just

4. What word begins like red?
   1. real  2. our  3. day  4. say

5. What word begins like did?
   1. black  2. dog  3. and  4. find
Objective 2.

Given specified words in a sentence the child will choose the correct ending sound for that word.

Directions: Select the correct letter from the list below the sentence and write it in the blank.

1. Mark and Janet like to swim_____.
   1. n  2. m  3. t  4. d

2. Ann will loo____ for the brown colt.
   1. b  2. k  3. t  4. l

3. Mother and Daddy went on a long tri_____.
   1. k  2. t  3. x  4. p

4. I see a big brown rabbi_____.
   1. t  2. p  3. k  4. v

5. I want to see a book about a far_____.
   1. b  2. d  3. n  4. m
Objective 3.

Given specified words in a sentence the child will choose the correct consonant sound for the medial position in the word.

Directions: Select the correct letter from the list below the sentence and write it in the blank.

1. Bill has a wa_____on.
   1. p   2. g   3. p   4. q

2. The ca_____el can eat grass.
   1. n   2. d   3. m   4. n

3. My coo_____ie is very big.
   1. k   2. r   3. t   4. z

4. I can make a se_____en.
   1. w   2. v   3. w   4. s

5. Here is a big red tu_____ip.
   1. b   2. d   3. h   4. l
Objective 4.

Given a written list of digraphs and an oral word beginning with one of the digraphs, the student will identify the digraph with the same sound.

Directions: The teacher will pronounce the word and the child is to circle the correct digraph.

1. Circle the digraph which begins like the word chirp.
   1. sh
   2. th
   3. ch
   4. wh

2. Circle the digraph which begins like the word shoe.
   1. sh
   2. th
   3. ch
   4. wh

3. Circle the digraph which begins like the word the.
   1. sh
   2. th
   3. ch
   4. wh

4. Circle the digraph which begins like the word whale.
   1. sh
   2. th
   3. ch
   4. wh
Objective 5.

Given a list of four words, the child is to circle the one word which has a different digraph sound.

Directions: Look at the words in each set. Circle the one word in each set which has a different beginning digraph.

1. then this white the

2. what white children which

3. she them shoe shadow

4. shadow church children chirp
Objective 6.

Given specified words in sentences and a choice of four blends selected from the following group; bl, st, sw, gr, br, sp, fl, br, fr, qu, and dr, the child will choose the correct blend.

Directions: Circle the blend which correctly completes each word.

1. Janet's ____ue coat was in the box.
   1. fr  2. tr  3. bl  4. dr

2. Ann ____oke the big doll.
   1. gl  2. bl  3. br  4. dr

3. Do you see the big black ____y?
   1. fl  2. st  3. cr  4. dr

4. John likes to ____im in the lake.
   1. st  2. gr  3. sp  4. sw

5. ____apes are good to eat.

6. You must ____op on red.
   1. bl  2. qu  3. st  4. sp

7. ____ot is a big dog.

8. My mother is a ____een.
   1. qu  2. sc  3. dr  4. br

9. A ____own is not a happy face.
   1. sc  2. dr  3. br  4. fr

10. I like to play a ____um.
    1. bl  2. dr  3. br  4. gr
Objective 7.

Given incomplete sentences and the alternate forms of a verb, ed, s, and ing, the child will select the proper form.

Directions: Circle the word that correctly fits in the sentence.

1. We like to go for _______ in the park.
   1. walks 2. walking 3. walked 4. walk

2. The puppy is _______ at that big dog.
   1. barked 2. bark 3. barking 4. barks

3. The bear is _______ on three legs.
   1. walked 2. walking 3. walk 4. walks

   1. work 2. works 3. worked 4. working

5. David _______ very well.
   1. singing 2. sing 3. singer 4. sings

6. Janet has _______ rope many times.
   1. jump 2. jumped 3. jumps 4. jumping

7. Mark _______ with his ball all of the time.
   1. play 2. playing 3. plays 4. play

8. I have not _______ that game.
   1. played 2. play 3. plays 4. playing

9. She _______ to school in the morning.
   1. walked 2. walk 3. walks 4. walking

10. I am _______ Janet to play.
    1. call 2. calls 3. called 4. calling
Objective 8.

Given an incomplete sentence the child will select the correct possessive form of the noun to complete the sentence.

Directions: Circle the word that correctly completes each sentence.

1. _______dress was very pretty.

2. The _______eyes are big and yellow.
   1. cat's  2. cat  3. dog  4. tiger

3. I am looking at _______old coat.

4. My _______car is blue.
   1. Mother  2. Daddy  3. son  4. Daddy's

5. The _______coat was very white.
   1. mouse  2. dog  3. dog's  4. cat
Objective 9.

Given incomplete sentences the child will complete the sentence with the correct plural form of the noun.

Directions: Circle the word that best completes each sentence.

1. See the three red ________.
   1. ball  2. balls  3. Ball  4. ball's

2. The two ________ are on the grass.

3. They will play some ________.
   1. games  2. game  3. game's  4. Game

4. How many ________ make a dime?
   1. nickle  2. nickle's  3. Nickle  4. nickles

5. Three ________ are jumping up and down.
   1. boy  2. boys  3. Boy  4. boy's
Objective 10.

Given a sentence with an incomplete form of a compound word, the child selects the corresponding part of the word.

Directions: Circle the word that best completes each part of the compound word.

1. Green grass grows in the spring ________.
   1. time  2. down  3. way  4. day

2. My birth ________ is in May.
   1. time  2. down  3. day  4. hill

3. The boy went uphill and down ________.
   1. way  2. time  3. hill  4. thing

4. The ________ born calf was brown.
   1. time  2. way  3. new  4. day

5. You may jump up your ________.
   1. bird  2. self  3. way  4. day
Objective 11.

Given one sample word and a set of four other words, the child will select the word that has the same phonetic part, ew (new), oo (too), ow (how), and oo (good).

Directions: Circle the word that has the same phonetic part as the sample word.

new 1. now know few well
too 2. zoo book many more
how 3. few cow blue what
good 4. may door so took
too 5. look moon two let
new 6. show dew foot shadow
too 7. foot book soon may
good 8. too cook new cake
how 9. stop plow look few
good 10. shook two what cow
Objective 12.

Given one sample word and a list of four other words the child will be able to identify the words that have the same vowel sound as the sample word.

Directions: Circle the word that has the same vowel sound as the sample word.

1. What word has the same vowel sound as the word man?
   1. make  2. fox  3. pan  4. came

2. What word has the same vowel sound as the word pin?
   1. time  2. fin  3. top  4. box

3. What word has the same vowel sound as the word bed?
   1. Ted  2. sleep  3. sheep  4. top

4. What word has the same vowel sound as the word fox?
   1. make  2. pop  3. cold  4. cat

5. What word has the same vowel sound as the word cup?
   1. cat  2. cute  3. fun  4. coat

6. What word has the same vowel sound as the word came?
   1. make  2. mat  3. moon  4. mit

7. What word has the same vowel sound as the word mine?
   1. came  2. time  3. mat  4. pat

8. What word has the same vowel sound as the word sheep?
   1. box  2. web  3. teeth  4. gun

9. What word has the same vowel sound as the word coat?
   1. boat  2. cot  3. box  4. like

10. What word has the same vowel sound as the word mule?
    1. music  2. mut  3. fin  4. pin