
7-1-1987

Evaluation/Testing Procedures In Reading

Elene S. Demos
University of Alberta

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons



Part of the Education Commons

Recommended Citation

Demos, E. S. (1987). Evaluation/Testing Procedures In Reading. *Reading Horizons: A Journal of Literacy and Language Arts*, 27(4). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol27/iss4/4

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons: A Journal of Literacy and Language Arts by an authorized editor of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



EVALUATION/TESTING PROCEDURES IN READING

ELENE S. DEMOS
Edinboro University
Edinboro, Pennsylvania

Introduction

This is the second in a series of four articles devoted to parents and reading. The first article described typical school reading programs and illustrated the key ingredients of each. This article treats the evaluation and testing procedures that schools use in the area of reading.

Types of Reading Tests

Schools have available a number and variety of well constructed tests that can be used in reading. Reading-Achievement/Survey Tests, Diagnostic Tests and Informal Reading Inventories are available for teachers. Each type of test will be examined in this paper.

Reading Achievement Tests

Reading achievement or survey tests are tests which give an approximate level at which a child functions in reading. They present a very broad picture of how a child is performing in comparison to others. Tests such as the California Achievement Tests: Reading, the Iowa Test of Basic Skills, the Stanford Achievement Test and the Metropolitan Achievement Tests: Reading, are examples of widely used group reading achievement tests. These tests may measure somewhat different skills and abilities, but each provides a general reading achievement score. On each test, at least two areas are measured--word recognition and comprehension.

Frequently the word recognition section of a test might be titled: --vocabulary

--word meanings

--word attack

--word knowledge

Tests differ as to how they ask students questions. For example, one vocabulary test may ask a student to select a word that is the opposite of a key stimulus word:

Sample A: on | start go off
hot | in water cold

Another test may ask a student to find the word that a picture names:

Sample B
(picture of cat) five six two cat
(picture of ball) ball cat broom toad

A third type of task might include asking students to complete a sentence that contains a missing word:

Sample C
The dog ran after the _____. (cat, run)
Billy's _____ was painted red. (house, dollar)

Still another type of test requires students to select the word the teacher pronounces for them:

Sample D
hat ham house
(teacher reads hat and student circles hat)

The comprehension section of a test usually asks students to read one or several paragraphs and answer questions about what they have read. As children progress from early elementary grades (K-3) to later elementary grades (4-6) the tests include fewer pictures, longer passages to read and more specific areas are tested.

These tests are norm-referenced and allow you and the teacher to compare your child's performance to other children. This simply means that the publishers of the test sampled a wide-variety of reading skills, have tested many children from various parts of the country at different age/grade levels and at different socio-economic levels, have developed and piloted the tests using specific directions and time limits and do not allow for teacher assistance during the testing situation. Reading achievement or survey tests are frequently administered to a whole class of students. They help you and the teacher answer several questions:

1. What is the general reading performance of my child?
2. What are the general strengths and weaknesses of my child in reading?
3. Has my child made progress in reading during the year?

Diagnostic Reading Tests

A second major category of tests in reading include

diagnostic tests. These tests may be group or individually administered. A teacher or parent could request such a test be given if the child performs poorly on an achievement-survey test or if specific problems have been observed while the child is reading. Tests such as the Gray Oral Reading Test, the Gates McKillop Horowitz Reading Diagnostic Test, the Woodcock Reading Mastery Tests and the Durrell Analysis of Reading Difficulty are examples of individual diagnostic tests. These tests take from 15 minutes to 90 minutes of individual testing time and require much skill in test administration and analysis. Group diagnostic tests are also available. These may be given to several students and include such tests as the Doren Diagnostic Reading Test, the Prescriptive Reading Inventory and the California Phonics Survey.

These tests provide more specific information in several areas of reading and typically include such areas as phonics skills, blending skills, visual memory, structural analysis, word recognition, spelling, syllabication, handwriting, knowledge of letters, names and sounds as well as comprehension.

These tests are used to help answer the following questions:

1. Does my child have specific skill weaknesses in reading?
2. Is my child's problem one of knowledge of basic skills or application of the skills?
3. What are my child's strengths in reading skills?

Diagnostic tests will help the teacher pinpoint specific strengths and weaknesses and provide a more comprehensive picture of the child's performance in reading.

Frequently, diagnostic tests are criterion-referenced. These tests are usually very narrow in focus and center upon specific reading skills. Their purpose is to measure which students show mastery of specific skills. These tests are frequently used throughout the school year and provide a monitoring device for student progress.

Informal Reading Inventories

Another valuable tool available for classroom teachers for diagnosing reading strengths and weaknesses include informal reading inventories (IRIs). Tests such as the Basic Reading Inventory, Classroom Reading Inventory, and the Analytical Reading Inventory. This type of test may be commercially

developed or teacher-made. Contrasted with the more formal standardized tests, the IRIs do not have standardized norms or formalized procedures.

They are intended to be used by the classroom teacher, a reading specialist or a diagnostician to examine a student's patterns of errors and to help the teacher determine different reading levels on each child, reading fluency and comprehension. For example, most IRIs provide the teacher with information about the following reading levels--the independent reading level, the instructional reading level, the frustration level and the listening level.

It is important for you to know these varied levels and what they mean. Briefly, the independent level is the level at which a child can read with ease. Books at this level present the student with few problems and are associated with pleasure reading material. At the instructional level, a student may encounter some difficulty, but with the instruction that the teacher provides, the child can continue to learn and read. The frustration level is too hard for the child and materials at that level create tension and anxiety for the student, the student miscalls many words and cannot understand what is read. In short, this level is too difficult for the child and should not be used with him. The listening level is frequently called the capacity or potential for reading. At this level, students can understand what is read to them. So even though they cannot read the words on the page, they can comprehend what they have heard.

This type of test frequently contains a graded word list, and several sets of reading passages with comprehension questions. These passages are read orally and silently by the child and some are read to the child. After the reading, children are asked questions by the teacher.

In addition to providing the teacher with information on the previously mentioned, the teacher can analyze the errors that children make while reading orally and answering questions. For example, if a student reads hat for bat, how for cow, sold for told, or not for hot, our analysis could indicate that a student has difficulty with beginnings of words and this might mean that he does not know how to apply sounds at the beginning of words, or that he may be in a hurry and not looking carefully at the words or not using the surrounding words to help him understand.

Further analysis is frequently done with the comprehension questions. The questions are written at different levels of difficulty. For example: When did Columbus discover America? is a very low level question and merely tests whether a child can remember and restate a date. A question such as--How do you think our astronauts are similar to Columbus and his crew? is a far more complex, higher-order question. Informal reading inventories allow teachers to determine the problem areas in comprehension and to plan a program to help students with their weaknesses.

The informal reading inventory is used to help answer the following questions:

1. What are my child's independent, instructional, and frustration levels?
2. What types of questions can my child comprehend?
3. How does my child correct his/her errors?

Test Results and Interpretations

After understanding the different types and purposes of reading tests, the next area to be considered is the result or scores. What do these scores mean? What is a raw score, a grade equivalent score, a percentile score, a percentage score, or a stanine? These terms are frequently used when discussing test results but are frequently misunderstood.

Let's examine achievement or survey test results first. The results of these tests are frequently used to evaluate where your child scores in reading in comparison to other children in the class, the school or the nation. The scores are frequently presented in different ways that include grade equivalent units, percentile scores or stanines. Frequently, teachers will refer to the raw score. The raw score is the number of items a student gets correct on a test. For example, if a student gets 48 items correct out of 50, the raw score is 48. A grade equivalent unit represents the score as a grade level score. A score of 3.2 indicates the child is reading at the grade equivalent level of the second month of the third grade, likewise, a score of 4.8 indicates a student is reading at the equivalent of the eighth month of the fourth grade.

Many achievement tests use the percentile score. A percentile of 86 means that your child did as well as 86

percent of the children, or could be interpreted to mean that 14 percent of the students performed at a higher level than your child. The percentile score does not mean that your child had 86% of the answers correct. This is frequently misunderstood by parents and educators. A stanine score also reflects a child's rank with others. Tests are divided into 9 stanines, a combination of two words - standard and nine. There are nine levels on a scale, the lowest score being 1. The middle stanine then, is 5 and indicates average performance. If a child scored in stanine 9 s/he would be performing at a very high level in which only about 4% of the population fell. A stanine of 1 would compare to the low 4% on the scale. The word standard in stanine is based on the fact that each stanine level is one-half of a standard deviation wide (except for 1 and 9).

Figure 1
Comparison of Stanine & Percentile Scores

Stanine	1	equals	percentile	0 - 3
"	2	"	"	4 - 10
"	3	"	"	11 - 22
"	4	"	"	23 - 39
"	5	"	"	40 - 59
"	6	"	"	60 - 76
"	7	"	"	77 - 88
"	8	"	"	89 - 95
"	9	"	"	96 - 99

Figure 1 illustrates the relationship between stanine and percentile scores. As you can see, stanine 5 or percentile 40-59 is at the midpoint and would indicate average performance.

Another important concept to remember about test scores is that some measurement error occurs in all tests. There are no absolutes in testing and we must consider ranges of scores. There is not enough precision in testing and tests; to pretend any different would be misleading. Even though the tests are prepared by experts who may use the best procedures for writing, selecting, and sampling items, some error occurs and it is best to think in terms of ranges.

Some Diagnostic Tests described previously, are norm referenced tests, but many are criterion-referenced and

report a percentage score. For example, if there were ten items on a test and your child's raw score (number correct) was eight, the percentage would be 80% correct. For many criterion-referenced tests, scores of 70-100% are considered to be mastery level.

The results from the informal reading inventory are presented as levels. The four levels are the independent, instructional, frustration, and listening. These levels help parents and teachers make decisions about the appropriateness and selection of materials for different reading purposes. For example, one wants to help children select easier reading material for pleasurable reading. More difficult material may be used when the teacher is working with a child in the school setting. Teachers and parents must also know when some material is too difficult for their child, so to prevent frustration, tension and anxiety, they would divert children from reading material which is too difficult.

SUMMARY

The three major categories of tests are achievement/survey, diagnostic and informal reading inventories. The purposes of each and the means of reporting scores are different for each type of test. These tests allow you and your child's teacher to monitor his/her progress, make instructional decisions and make comparisons between children.