DOCUMENT RESUME

ED 043 452 RE 002 931

AUTHOR Winkley, Carol K.
TITLE What Do Diagnostic Reading Tests Really Diagnose?
PUB DATE 7 May 70
NOTE 21p.; Paper presented at the International Reading Association conference, Anaheim, Cal., May 6-9, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.15
DESCRIPTORS Achievement Tests, Reading Comprehension, *Reading Diagnosis, Reading Level, *Reading Skills, *Reading Tests, *Standardized Tests, Test Interpretation, Test Selection, Verbal Stimuli, Visual Stimuli, Word Recognition, Word Study Skills

ABSTRACT

A study was made of nine reading tests, including both group and individually-administered measures, which are claimed to be chiefly diagnostic. Instruments analyzed were the following: Silent Reading Diagnostic Tests (Bond, Balow, and Hoyt), Botel Reading Inventory, Durrell Analysis of Reading Difficulty, Gates-McKillop Reading Diagnostic Tests, McCullough Word Analysis Tests, Roswell-Chall Diagnostic Reading Tests, Diagnostic Reading Scales (Spache), and Levels I and II of the Stanford Diagnostic Reading Test. Examination of the nine diagnostic batteries revealed subtests for (1) measuring potential reading level, (2) measuring silent and oral reading performance, (3) estimating reading levels, (4) identifying inhibiting factors, (5) determining chief skill deficiency area, (6) determining word identification technique, and (7) locating word recognition difficulties. Among the conclusions, it was stated that most instruments cannot be used to determine the chief area of skill deficiency, including specific problems of vocabulary, comprehension, and rate. Word recognition subtests are limited in scope of subskills assessed and emphasize spelling ability. In addition, skills for monosyllabic words are more often measured than skills required to unlock polysyllabic words. Group-administered tests are limited to silent activities. (WB)

Carol K. Winkley
Associate Professor of Education
Northern Illinois University
DeKalb, Illinois 60115

What Do Diagnostic Reading Tests Really Diagnose?*

Session: Evaluating Reading Achievement

Thursday, May 7, 1970 10:45 - 11:45 a.m.

Introduction

In response to the question, "What is a good diagnostic reading test?" the answer, "There is none!" is frequently given. If the inquirer is seeking a single instrument equally appropriate for all levels and suitable for locating problems in all skill areas, the response is, no doubt, a valid one. Yet, there is a need for instruments that classroom teachers can use to supplement their judgments based on diagnostic teaching, and that clinicians can use to pinpoint problem areas when conducting a clinical diagnosis. What is available?

*The author wishes to acknowledge the assistance of Benita Vyverberg, graduate assistant, in collecting the data for this study.

In order to answer this question, a study was made of nine reading tests, including those planned for both individual and group administration, which are claimed to be chiefly diagnostic instruments. Titles of the tests examined are found at the top of the columns in Table 1.

(Table 1. pp. 1 & 2)

The first entry across the table provides information regarding the grade and/or reading levels for which the test was intended. The second entry indicates whether the test must be given individually or whether it can be used in a group situation. A careful analysis of each subtest and its stated or implied purpose revealed that these nine instruments contained subtests for:

1. measuring potential reading level
2. measuring silent and oral reading performance
3. estimating independent and instructional reading levels
4. identifying inhibiting factors
5. determining chief area of skill deficiency
6. determining technique of word identification
7. locating work recognition difficulties

It is clear that no common definition of a diagnostic reading test is held by the authors of these tests, nor have the various authors had similar purposes in mind as they developed their tests.

Assessments of Potential Reading Level

Six instruments contain subtests purporting to estimate the child's potential level of reading achievement. Three general types of activities were utilized by the various authors:

1. Listening comprehension of paragraphs read aloud by the teacher (D) (Sp);
2. Selecting appropriate meanings of words presented orally (GM) (St I) (St II); and

3. Selecting word opposites as words are read aloud (B)

No doubt the subtests described are included in these diagnostic batteries to enable the teacher to easily determine whether each child is disabled in reading (reading at a level significantly below his ability level).

Measures of Silent and Oral Reading

Two instruments (GM) (Sp) provide a subtest of "Oral Reading" while the Durrell Analysis of Reading Difficulty includes a subtest of "Silent Reading" in addition to "Oral Reading." In each instance, successive paragraphs, each increasing in difficulty over the previous one are read. As a measure of unaided recall, the child retells each story to the examiner in the "Silent Reading" section of Durrell's test. A simple comprehension check follows each paragraph, except in the Gates-McKillop battery.

These subtests appear to have three possible purposes: (1) to provide an opportunity to record and analyze types of oral reading errors; (2) to make it possible to compare difficulties in silent reading with those in oral reading, as in Durrell's test; and (3) to make it possible to compare reading achievement with some measure of ability in order to determine whether a child is truly disabled in reading.

Estimates of Independent and Instructional Levels

Recognizing the necessity for teachers to be able to locate and then provide instruction at the appropriate levels, both Botel and Spache have included subtests in their diagnostic batteries which they suggest be used to estimate independent and instructional levels.

Each author includes graded word lists and specific criteria to be applied in estimating reading levels. Spache also includes graded reading passages for the child to read aloud while Botel incorporates a "Word Opposites Reading Test" into his battery. The remaining authors of diagnostic reading tests made no provision for determining the actual reading levels of the pupils taking their tests.

Identifiers of Inhibiting Factors

Inhibiting factors are those characteristics of the child and/or the home and school environment which are preventing normal progress in reading. Correcting or alleviating them will make it possible for the child to learn to read with greater ease but the reading difficulty itself must still be identified and skill deficiencies eliminated through remedial teaching.

Both the Durrell Analysis of Reading Difficulty and the Gates-McKillop Reading Diagnostic Tests contain subtests of certain visual and/or auditory aptitudes. The administration of three subtests of the Durrell battery makes it possible to determine whether a child has strengths or weaknesses in both visual and auditory skills. At the primary level the subtest "Visual Memory of Words" provides an evaluation of a child's ability to select a word, seen in a brief exposure using a tachistoscope, from several words of similar configuration. "Hearing Sounds in Words" is a subtest requiring the child to select the word printed in the test booklet that begins, ends, or begins and ends with the same sound(s) heard in words pronounced by the examiner. An analysis of the child's errors on a

third subtest, the "Spelling Test," will often disclose additional information concerning the child's relative use of visual memory and phonic principles in writing words having both regular and irregular spelling.

Three similar subtests, which are much more difficult, are provided for students reading at the intermediate grade levels: (1) "Visual Memory of Words--Intermediate" which requires the child to write the word seen in a brief tachistoscopic exposure; (2) "Phonic Spelling of Words" in which the child is asked to spell words just as they sound. (Credit is given for any type of phonetic spelling); and (3) "Spelling Test." In order to identify visual and/or auditory strengths and weaknesses, the same types of comparisons can be made as those suggested at the primary level.

In the Gates-McKillop battery, three subtests of auditory skills are found: (1) Auditory Discrimination; (2) Auditory Blending; and (3) Spelling. With a sample of fourteen items, the child is asked to tell whether two words pronounced by the examiner are the same or different words. As an indication of a child's auditory blending ability, he is asked to pronounce a word as a whole, which he has heard the examiner say part by part. On the spelling test the words are spelled aloud by the child to enable the examiner to determine whether the child spells letter by letter or by phonic elements.

"Auditory Discrimination" is also a subtest of the Stanford Diagnostic Reading Test, Level 1. The format of this test differs from that of the two tests of auditory discrimination mentioned earlier, thus making it possible to administer the test to groups. After the

teacher pronounces two words, the child makes an X through B in his test booklet if the words begin the same, through E if they end the same, and through M if the middle sounds are the same.

Determiners of Chief Area of Skill Deficiency

In making a diagnosis of a child's reading difficulties, per se, the diagnostician's first task is to determine the chief area of skill deficiency as being in word recognition, vocabulary or word meanings, quality of comprehension, or rate of comprehension. It is estimated that 90 to 95 percent of the children who have trouble with reading have deficiencies in the area of word recognition which in turn affect obtaining the meanings of the words, understanding what is read, or the speed of reading. This means that 5 to 10 percent of disabled readers will not have any major problems in the area of word recognition and can be expected to have as their chief area of weakness either vocabulary, comprehension, or rate. Do diagnostic reading tests help to determine a child's chief area of skill deficiency?

There are subtests in four of the diagnostic instruments examined which will give the diagnostician some help. The remaining test batteries have subtests in one skill area only--word recognition. Each of the four instruments mentioned above includes some measure of vocabulary but only the Stanford Diagnostic Reading Tests include subtests of comprehension and rate.

The Stanford tests and the Gates-McKillop Reading Diagnostic Tests each include a listening test of vocabulary described earlier as a measure of potential reading level. The child is required to do no reading, thus providing information regarding his knowledge of

word meanings which is not hampered by inability to attack unknown words met in silent reading. If the grade score appears to be low in comparison with the child's performance on various word recognition subtests, there may be evidence that vocabulary should be considered the child's chief problem area.

The "Word Opposites Tests" (Reading and Listening) of the Botel Reading Inventory are not so much measures of comprehension, as the author states in his manual, as they are tests of vocabulary or knowledge of word meanings. He suggests that a comparison of scores obtained when the test is read silently with scores earned when the teacher reads the words aloud would help to identify those pupils whose reading performances were significantly lower than their reading potentials. However, these tests can serve another purpose. If the listening score is considerably higher than the reading score, the child can be suspected to be weak in word recognition rather than vocabulary.

In the "Reading Comprehension" subtests of the Stanford Diagnostic Reading Tests, numbered blanks appear in paragraphs which the child reads silently selecting from four choices the word that belongs in each space. Level 2 contains a subtest of "Rate of Reading" in which children are timed in reading content of uniform difficulty and selecting an appropriate word from three choices in every third line to fit the meaning of the sentence. After the raw scores have been converted to stanines, the stanine ratings can be compared. A difference of two or more stanines between subtests is indicative of a possible area of skill deficiency.

"Diagnosers" of Difficulties in Vocabulary, Comprehension and Rate

Compared with word recognition, few disabled readers have major

difficulties in the areas of vocabulary, comprehension and rate. Those students who do have trouble in these areas are most frequently found at the junior and senior high school levels. Nevertheless, instruments to determine whether a child's vocabulary difficulties are due to a lack of understanding of prefix and suffix meanings, not knowing multiple meanings of words, lack of dictionary skills, etc., are needed. There are none. The same situation exists when one looks for a diagnostic test of "Rate." There is no test to determine a child's flexibility of speed when reading for different purposes, for example.

The only diagnostic reading test to provide a breakdown of the child's comprehension skills was the Stanford Diagnostic Reading Test, Level II. About half of the items test literal comprehension and the remaining items check inferential comprehension. When one comprehension score is two or more stanines below the other score the child may need remedial instruction in that area of comprehension. The examiner still will not know whether the child needs help in understanding main ideas of selections, in understanding sequence, or in recalling facts or details, for example, even though he has been found weak in literal comprehension. Neither are subskill deficiencies identified in the broad area of inferential comprehension.

Determiners of Technique of Word Identification

All of the diagnostic instruments that were examined contained subtests, listed in Table 1, which assessed a variety of word recognition skills. Not all of the tests included subtests of instantaneous word recognition which could then be compared with another subtest in which the child was given sufficient time to use his phonic and structural analysis skills to attack the words not recognized at sight.

Such a comparison makes it possible for the examiner to determine not only the size of each child's sight vocabulary but the extent to which he can use various word recognition skills. If the child's knowledge of phonic and structural analysis skills has not developed to a point where he can use them in attacking unknown words, he has not yet acquired them and needs further instruction.

Subtests providing the opportunity to compare flash presentations of words with untimed presentations are found in three diagnostic batteries. (D) (GM) (Sp) A hand tachistoscope is used in the first two instruments, whereas the examiner merely checks words that a child recognizes instantaneously as he reads lists of words in Spache's "Word Recognition" subtest. In each instance the child is given more time to carefully analyze any word not recognized at sight.

Locators of Phonic Problems

All of the diagnostic reading tests analyzed contained three or more subtests of word recognition skills. (See Table 1.) In order to evaluate each test battery as it would function in the identification of a child's chief skill deficiencies in word recognition two steps were taken. First, the pretest steps suggested in a previous paper* and listed in Table 2 were used to determine the level of understand-

(Table 2)

ing required of a child to perform successfully on any subtest. Second, each subtest of phonic skills was then categorized at the pretest level that most nearly approximated the behavior expected of the testee.

An examination of Table 2 reveals that no subtests were categorized at the lowest three pretest steps. Examples of such behaviors are often

*Winkley, Carol K., "Why Not an Intensive-Gradual Phonic Approach," Reading Teacher, (Vol. 23, No. 7), April, 1970, pp. 611-617, 620.

called for as a part of a readiness evaluation. Subtests requiring the naming of capital and/or lower case letters are found on the Durrell Analysis of Reading Difficulty and the Gates-McKillop Reading Diagnostic Tests.

Selecting the written representation, or grapheme (from a group of four or five letters) corresponding to a sound heard in a word pronounced by the examiner is a common response required on group instruments. However, four subtests of the Gates-McKillop Reading Diagnostic Tests, an individual battery, called for a similar behavior. An interesting variation of this technique of requiring the child to match grapheme to phoneme is found in the "Blending" subtest of the Stanford Diagnostic Reading Test, Level 1. The teacher pronounces a word, such as "trick", for which the child is to select the appropriate beginning, middle, and ending from two choices for each.

Example:	⊗ tr	⊗ i	○ ch
	○ br	○ e	⊗ ck

Pretest Step 6 is only a slight variation of Step 5 requiring the pupil to select a printed word in which the letter appears, instead of a single grapheme, that stands for a particular phoneme (or phonemes) heard in a word pronounced by the diagnostician. A subtest of this type, "Hearing Sounds in Words", appears in the Durrell Analysis of Reading Difficulty. The "Phonetic Discrimination" subtest of the McCullough Word Analysis Tests differs slightly because the pupil must identify the stimulus word himself, such as "blow" and then find the word among four choices in which he hears the sound of the underlined letters.

Example: out not horse old

The absence of an auditory stimulus increases the difficulty of this exercise. A subtest in the Gates-McKillop battery is somewhat different, also, because the child is directed to select from four nonsense words printed in his test booklet, the one pronounced by his teacher.. For example, the teacher might say, "spə nĕs'". These spellings appear in the test booklet: spiness stinacc spiss squents

At a higher level pretest step, where pictures supplant the auditory stimuli, the child is asked to find the grapheme(s) representing the sound(s) heard in the name of a pictured object. A subtest of the Stanford Diagnostic Reading Test, Level 1, utilizes this technique. The child selects the two- or three- letter combination standing for the sounds heard at the beginning or the end of the word represented by the picture. The subtest entitled "Words in Isolation" of the Silent Reading Diagnostic Tests calls for selecting an entire word to go with a picture. The test differs from the ordinary vocabulary test at the primary level in that the foils are not all real words but represent beginning, ending, middle, or orientation errors that a pupil might make. The key for scoring is coded to enable the teacher to classify the types of incorrect choices made by each child.

Like Pretest Steps 5, 6, and 7, Pretest Step 8 is more closely related to spelling than reading. The child is required to recall the grapheme representing a phoneme heard in a word, which is a spelling skill--not a reading skill. Only to the extent that word pronunciation and spelling are related can these tests be considered valid measures of a child's use of phonics in pronouncing unknown words.

In the Botel Reading Inventory, the child writes the grapheme

representing the phoneme heard at the beginning, end, or middle of a spoken word. Spache, in his Diagnostic Reading Scales, has the pupil write the letter representing isolated phonemes sounded by the teacher. At the intermediate grade level the Durrell Analysis of Reading Difficulty has a subtest requiring the students to write phonetically certain words not normally appearing in their vocabularies, such as "carpolite." Any phonetic spelling is judged correct, even "karpulight."

Giving the sound represented by a separate letter (Step 9) tests a skill needed in reading, and yet adequate performance on this level does not insure the child's ability to blend the sounds and accurately pronounce an unfamiliar word. Since this is an ability that must be checked individually, only the tests developed for individual administration include subtests requiring this behavior of the testees. (D) (GM) (RC) (Sp) See Table 2.

It makes sense that if a diagnostician wants to find out how well a child uses phonic skills to pronounce an unknown word, he should give the child some unknown words to pronounce. How can he be sure he has selected unknown words? One way is to use nonsense words like those found in subtests of the Gates-McKillop Reading Diagnostic Tests and the Botel Reading Inventory. Several subtests of the Roswell-Chall Diagnostic Reading Tests use real words that are not normally in the sight vocabulary of a child at the lower levels who is having difficulty with reading. Spache has two subtests requiring the child to pronounce groups of letters: (1) "Vowel Sounds", which has several four-letter words, each containing a different vowel letter, to be pronounced first with the long sound of the vowel and then the short sound; and (2) "Common Syllables", many of which are phonograms, to be pronounced in isolation.

Since any technique requiring pupils to respond verbally can not be incorporated into a group instrument, the authors of two group tests have developed subtests which come close to requiring the same behavior of the children taking the test. McCullough in her group test includes a subtest "Sounding Whole Words" in which the child must select a word from three unfamiliar groupings of letters by sounding each phonetically. As a fourth option he may put a cross in a blank if no word in the row sounds like a word he knows.

Example: spayss trayk smay ____

In the "Blending" subtest of Level II of the Stanford tests, a format similar to that in Level I is used. However, at this higher level the teacher does not pronounce each word but the child must sound the elements and blend them together to be sure he has put together a meaningful word.

Locators of Difficulties in Structural Analysis

Compared with the number of subtests found in diagnostic instruments that evaluate various levels of a child's phonic knowledge, there are relatively few tests of structural analysis skills. These have been categorized in Table 3 under: (1) Locating Root Word; (2) Syllabication; (3) Blending; and (4) Accent. Each subtest was

(Table 3)

examined to determine whether an auditory or visual stimulus was presented, and the response required was ascertained as oral or written.

Two subtests involve locating the root word in an affixed word. In both instances the child provides a written response to a visual stimulus appearing in group instruments. (BBH) (Mc) In the recent

test developed by Bond, Balow and Hoyt, the child is asked to select the root word, among three choices, from which the word appearing in the first column was made ("Visual - Structural Analysis"). However, in "Root Words in Affixed Forms" (Mc) the children are directed to circle each prefix and suffix. In some of the words, however, the part remaining when the so-called prefix is circled is not a root word. For example, "mend" is not the root word of "commend"; nor does "invite" have "vite" as its root word. Similar errors are noted in the Bond, Balow, and Hoyt subtest.

Subtests of syllabication skills, found in seven of the nine instruments, were classified on three different stimulus-response levels. A variety of behaviors is expected of children. They are:

(1) circling a number to show the correct number of syllables in each word pronounced by the examiner (B);

(2) selecting from three choices the correct syllabic division of a word listed in the first column (BBH);

(3) drawing a line to separate the two syllables of a word (Mc);

(4) selecting the first syllable of words with one or more syllables (St I) (St II);

(5) reading multisyllabic words including compound words, affixed words, and words with inflectional endings (RC); and

(6) reading nonsense words of two or more syllables (B) (GM).

To determine pupils' ability to blend the syllables and pronounce a word as a whole, a similar range of types of activities appeared on the various instruments. In response to an auditory stimulus, where the teacher pronounces the various phonic elements in a word separately, the child is expected to respond by pronouncing the word as a whole. (GM) From this lowest stimulus-response level (GM), tests of increasing difficulty and complexity appear on other batteries. (See Table 3.)

At the highest level the child pronounces words showing that he can blend their parts (GM) (Sp) and also demonstrates this ability in pronouncing the nonsense words on the Botel Reading Inventory.

Only two subtests in Botel's inventory provide any measure of a child's ability to determine the accented syllable. In the first, the child circles the number that shows which syllable is accented in each word that he hears. In the "Nonsense Words" subtest the examiner can observe the child's ability to place the accent on the correct syllable when pronouncing an unknown word. (Although accent is considered a phonic skill affecting vowel sounds rather than word structure, the subtests appeared to lend themselves to the classification scheme used for structural analysis skills.)

Summary and Conclusions

This careful examination of subtests on nine different diagnostic test batteries has revealed that:

1. these instruments have a variety of purposes, several of which are not truly diagnostic in nature.
2. most of the instruments cannot be used to determine a child's chief area of skill deficiency. This can probably be done better with a survey silent reading test, anyway.
3. it is not possible to pinpoint specific problems in the areas of vocabulary, comprehension, or rate with these instruments.
4. although there are many subtests of word recognition skills, most of them really evaluate spelling ability rather than reading ability.
5. group-administered tests are limited to silent-type activities often requiring the child to listen and select or supply graphemic representations of phonemic elements.
6. no single test, group or individual, assesses all subskills of word recognition from knowledge of consonant sounds to ability to select the accented syllable in an unknown word.
7. skills required to unlock single syllable words are measured more frequently than those required to attack multisyllabic words.

8. certain errors exist, particularly in the selection of affixed words.

Before the decision is made to use any part of a diagnostic battery, the examiner should ask himself, "Is this test evaluating an ability not measured better by some other instrument specifically developed to determine intellectual capacity or reading level?" It appears that several authors of diagnostic instruments have attempted to be "all things to all people". Shouldn't a diagnostic reading test be one that diagnoses the reading problem itself? Shouldn't it help the diagnostician to find each child's strengths and weaknesses in reading skill development? Shouldn't a diagnostic instrument provide some indication of the level to which a child's acquisition of a specific skill has progressed? Can we be sure a child knows a phonic skill well enough to use it in reading when he demonstrates the ability to use it in a spelling activity?

These are all questions that must be answered by future authors of diagnostic tests. If teaching strategy is to be determined by a careful analysis of each child's performance, subtests of diagnostic instruments must be constructed to pinpoint the child's difficulties in the reading act itself.

List of Diagnostic Reading Tests

Bond, Balow, Hoyt - Silent Reading Diagnostic Tests. Chicago: Lyons and Carnahan, 1970.

Botel Reading Inventory. Chicago: Follett Publishing Company, 1966.

Durrell Analysis of Reading Difficulty. Chicago: World Book Company, 1955.

Gates-McKillop Reading Diagnostic Tests. New York: Teachers College Press, Columbia University, 1962.

McCullough Word-Analysis Tests. Boston: Ginn and Company, 1962.

Roswell-Chall Diagnostic Reading Test. New York: Essay Press, 1959.

Spache, George D., Diagnostic Reading Scales. Monterey, California: California Test Bureau, 1963.

Stanford Diagnostic Reading Test, Levels I and II. New York: Harcourt, Brace and World, Inc., 1966.