

1-1-1973

Reading Tests Don't Really Tell It Like It Is

George D. Masters
Bedford Public Schools, Bedford, New York

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



Part of the Education Commons

Recommended Citation

Masters, G. D. (1973). Reading Tests Don't Really Tell It Like It Is. *Reading Horizons: A Journal of Literacy and Language Arts*, 13 (2). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol13/iss2/2

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.

READING TESTS DON'T REALLY TELL IT LIKE IT IS

George D. Masters

BEDFORD PUBLIC SCHOOLS, BEDFORD, N.Y.

Reading, at one time the hallmark of the educated man, is now the bugbear of the modern school. In other ages it was recognized that there were some children who didn't want to read, never wanted to read, and never would want to read. Today, such a student is spurred, goaded, and hounded till he either learns to read for self preservation, drops out of school, or confronted with failure, lets the world see the only R's he has learned, the wrong R's—resistance, rejection, and rebellion.

Although it is true that one can get more out of life from reading, it is not true that he who does not read is a failure. It takes intelligence to drive a tractor-trailer, to type a letter, or to repair a faulty carbureter, but it is not necessary to have the reading level of a college graduate. And this is what the majority of schools and parents are desiring for their students and children. Teachers want students to read well because they are judged on the reading ability of their students. Students are the finished product of the schools; and, like other industries, the school business is interested in turning out a product that is acceptable. And parents. Well, they are parents and are naturally anxious that their children succeed. But it is very difficult for the schools to judge a product that is acceptable. The modern educator realizes this and together with the parent clings to the lodestar of the standardized test—the one seeming constant in a confused constellation of pupil personnel data.

A standardized test is much like a recipe in cooking. If a woman in the kitchen wants to try a new dish, e.g. chicken chow mein, she will need instructions that have proven successful in the past. If she follows these instructions, she is reasonably sure of creating a culinary delight. The instructions that she is following have been reproduced over and over again to insure an accurate recipe. She is told to use one-half tablespoon of salt, not one; one-fourth pound of butter, not one-half; or one-half pound of rice, not a pound. In the same way standardized tests are subjected to rigorous study and experimentation. The conditions of testing, the reading of directions, and the scoring are always the same. Thus the child that is tested in Chicago, San Francisco, or New York will receive the exact same questions and

directions and his answers will be judged in the same way. The child's score is then compared with the scores of a mythical-like tested norm. His success or failure is based on where he measures when placed against this yardstick.

Although many parents and some teachers think that all standardized tests are intended to measure intelligence or "IQ," these tests may be divided into several kinds designed for different purposes.

The first kind attempts to measure intelligence or aptitude for school work. It stresses those kinds of ability that have to do with success in school and is composed largely of questions or items not specifically taught in school. The usual intelligence test will give one score for verbal or linguistic ability and one for numerical or mathematical ability. Some will even provide scores for memory, reasoning, spatial differentiation, and the like.

A second kind of standardized test is the interest inventory. Interest inventories are not really tests, since there are no right or wrong answers. The questions present to the student a list of possible likes and dislikes. Organized so that each response can be made by choosing from two or more suggested answers, the replies are then scored for a variety of occupations or occupational fields. The results show the relative strength of the individual's interests in different vocations, or how his interests compare with those of persons successfully engaged in various occupations.

A third kind, closely related to the second, is the personality measure. It also has no right or wrong replies. It is geared so that the individual answers questions about his hopes and fears, likes and dislikes, his actions in varying situations, and so forth.

A fourth kind tries to measure achievement in school subjects or subject fields—arithmetic, language, spelling in the primary grades and English, physics, history and other areas of study in secondary grades. For the most part it attempts to measure the knowledge of important facts learned in a subject area. The reading test is generally considered to fall into this category, although it should be given a classification all its own. The other tests are measures of amounts of knowledge. A reading test is an attempt to measure a skill. It is more concerned not with *what* a student knows but with *how* a student reads.

At one time a noted educator sat down and tried to determine what made a good reader. He finally came up with a list of 83 different skills that must be acquired if one is to read well. This complex dissection of the reading process showed that there were 17 separate

skills involved in understanding word meaning; 20 related to word analysis; 13 bearing directly on comprehension and interpretation; 25 involved in study skills; and 8 related to oral reading. The reading process of the good reader may be compared to the breathing process of a healthy individual. Both are so natural that one fails to realize how complicated each is, and only when something is wrong does he become aware. Few standardized tests attempt to measure half of these skills, much less all of them.

The standardized reading test is usually divided into two parts. The first part consists of vocabulary words to be defined. The reader is usually provided with a list of four or five words from which he makes a choice. The second part consists of paragraphs to be read followed by questions based on the paragraphs. As with the vocabulary, the reader is here also given a choice of answers. The results of these two sections may be given in stanines, percentile ranks, or grade equivalents. Of the three the grade equivalent or grade score is used most often. A score of 5.5 means that the reader has earned a score equal to the average score earned by children in the fifth month of the fifth grade. But those who make use of these reading tests often forget the basic assumption of the test makers: in a group of undifferentiated youngsters, some must always score higher on a test than others in accordance with the natural law of individual differences. Thus, the designers of the California Achievement Test found that the greatest number of fifth graders taking the test during the fifth month of school had 25 correct answers. This number then became the 5.5 reading grade for all future test takers. Too often people assume that any pupil with such a score could not understand and appreciate a book composed for the ninth grade despite his social, physical, and intellectual maturity. The bell shape curve of individual differences is not to be construed as a vertical line of homogeneous abilities. That which makes for differences in height, weight, and color, also makes for differences in intellectual ability. Each child is an individual and should not be made to fit into any preconceived mold.

But it is based on these scores that the destinies of hundreds of children are decided. They are often used to separate the sheep from the goats. Those who do badly usually wind up in the oaf's class and need influential fathers to get them into better colleges. These tests should only be another factor in judging the intellectual ability of a child. By themselves, they do not really "tell it like it is."

There are only two reasons why people read—for pleasure and for

information. How much pleasure and how much information do students get from a reading test? A reading test attempts to measure the testee's normal reading ability by placing him in an abnormal reading situation.

A reading test usually consists of isolated passages. This use of isolated passages, so reminiscent of the commonplace books of Francis Bacon, is unsound because of its very disjointedness. There is no smooth continuity. Coupled with these isolated passages is the complaint common to many of the students to whom I have given reading tests: "They're so boring." One of the more widely used reading tests contains short paragraphs on seals, hurricanes, an Arabian fable, Venice, Bach, an ichthyologist named James L. B. Smith, and apple growing in New Hampshire. For some this may seem like admirable fare. But for many children, and particularly the urban child raised on the streets of the inner city, Bach, apple growing, and the discoveries of ichthyology do not "hit them where they live." It seems strange that—if on one hand publishers realize that today's children are different and require a different kind of reading material than the Tom, Dick, and Jane stories of yesteryear, as witness any one of the new reading series now being published—these same publishers should use such innocuous reading passages in those tests which measure the reading ability of children, many of whom lack the money or parental interest to remedy an unjust testing score.

Many of these tests are geared for the commonplace and mundane. They discriminate against the imaginative mind. One test, for example, asks the student to choose one of five words "that means the same, or almost nearly the same" as "quiet." The choices: "exact," "still," "tense," "watery," and "blue." The dull reader of course will reply "still," which is the test-approved answer. But the child who reads creatively and replies with the poetic "blue" fails. Too often the testers award the scholarship of the humdrum, and a high grade often indicates a safe student, one who may never have thought of thinking for himself at all.

The lesson of Aesop's tortoise appears to be lost on the modern test designers. In athletic events where the sprinter is measured by time to bring out his best effort, the use of minutes and seconds is quite valid. But in most reading tests where the testee is given an X amount of time to finish a section, the time factor works against the thoughtful, slow reader. The introduction of this irrelevant factor implies that the slow reader is less competent than the rapid.

In scoring, most tests also discriminate against the slower and more deliberate student. Few tests take into consideration the number

of correct responses in proportion to the number of questions attempted. The tortoise-like student will receive a lower score although in proportion to the number of questions attempted, he may have a very high percentage correct. In order to compensate for this, teachers sometimes instruct students to answer every question. They feel that according to the law of averages these students are bound to guess one or two answers correctly and thus increase their scores. Such encouragement seems to indicate that reflection and thoughtfulness are held in low esteem while glibness and shallowness are lauded. Sometimes, as one looks about, one is tempted to say that this is as it should be. The age of the thoughtful reader may be a thing of the past.

This encouragement by many reading tests of thoughtlessness and glibness is not limited to the students or test takers. The reading tests encourage the same in the test givers—the teachers. Most tests give results in stanines, or percentile ranks, or grade scores, as was mentioned earlier. All of these are judgments, and merely encourage a teacher to test, score, and file. If the test results offered an analysis of the skills measured, maybe the youngster who is weak in finding the main idea might be able to receive some individual instruction in this skill. If the test makers could provide some sort of a reading profile of the test taker through something like an item analysis, perhaps then test scores and results might be of more value.

A reading test does not measure reading skills. It does measure a pupil's exposure to the written word. Given the same intelligence, two students will differ in reading maturity according to the number of words they have read. The student who has read 200,000 words in short stories and novels is bound to score higher than the youngster who has read only 20,000. Both may be compared to the man who goes for a walk in the country with his dog. The man walks straight along; the dog dashes back and forth, investigating anything and everything that attracts his interest. He dashes into bushes, chases squirrels, and sniffs everywhere. Both man and dog reach the same destination, but only the dog has really explored the whole terrain. And the student who has toiled over 200,000 words will master vocabulary much more easily than the student who has glanced at only 20,000 words.

Test makers, test givers, test takers, and all others who make use of test scores might do well to remember the advice that Montaigne offered in one of his essays, "Man is a marvellous, vain, fickle, and unstable subject, and on whom it is very hard to form any certain or uniform judgment."