DOCUMENT RESUME

ED 047 911 RE 003 366

AUTHOR Carver, Ronald P.
TITLE The Criterion-Referenced Aspects of the Carver-Darby Chunked Reading Test.
PUB DATE Dec 70
NOTE 8p.; Paper presented at the National Reading Conference, St. Petersburg, Fla., Dec. 3-5, 1970
AVAILABLE FROM Twentieth Yearbook of the National Reading Conference, Inc., 1217 W. Wisconsin Ave., Milwaukee, Wis. 53233. In press

EDRS PRICE EDRS Price MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS *College Students, Measurement Instruments, *Performance Criteria, Reading Ability, *Reading Comprehension, *Reading Speed, *Reading Tests, Test Construction

ABSTRACT

In addition to using chunked items, the Carver-Darby Chunked Reading Test differs from the traditional reading test in the way it was developed and the manner in which the test scores are interpreted. The criterion for developing and revising test items was based on the inability of the readers to answer test items correctly before reading the passages and their ability to answer the test items correctly after reading the passages. Interpretation of the test depends on three scores: (1) the efficiency score, i.e., the number of incorrect chunks out of 100 possible that the individual correctly identifies during the 25-minute test period; (2) rate, which is determined by the number of the last item attempted; and (3) accuracy, which is the percent correct. Using these scores the reader can be classified as efficient-inefficient, rapid-slow, and accurate-inaccurate which yield six possible categories of mature readers. Although there is no empirical research which supports this particular categorization, speculative judgments are made concerning them with the hope of stimulating definitive research. References are included. (DH)

Paper presented at the meeting of the
National Reading Conference
St. Petersburg, December, 1970

The Criterion-Referenced Aspects of the Carver-Darby Chunked Reading Test

Ronald P. Carver
American Institutes for Research, Washington, D. C.

ABSTRACT

In addition to using chunked items, the Carver-Darby Chunked Reading Test also differs from the traditional reading test in the way it was developed and the way that test scores are interpreted. It was not developed to maximize individual differences and test scores are not compared to a norm-group set of scores in the traditional fashion. The test was developed so that a criterion group of readers could answer the items and a criterion group of nonreaders could not answer the items. Furthermore, the possible score combinations on the test were used to designate six types of readers. Hypothesized attributes of the six types of readers were presented in some detail. Research is needed to evaluate the typology.

The Criterion-Referenced Aspects of the
Carver-Darby Chunked Reading Test

Ronald P. Carver
American Institutes for Research, Washington D. C.

A chunked type of reading comprehension measure was suggested by Carver (1), and subsequently a standardized chunked test of information stored during reading has been developed, evaluated (Carver & Darby, 4), and analyzed (Carver & Darby, 5). Recently, this test has been published as the "Carver-Darby Chunked Reading Test" (Darby & Darver, 6). The advantages and disadvantages of the Carver-Darby Chunked Reading Test (CDCRT) have been discussed in relation to the traditional multiple-choice type of reading test and to the cloze type of test (Carver, 2). Not explicitly discussed in previous journal publications are the aspects of this test which are relevant to criterion-referenced measurement (See Carver, 3 or Popham & Husek, 7). The purpose of this presentation is to breifly discuss the development of the test and what it proposes to accomplish from the standpoint of criterion-referenced measurement.

Traditionally, standardized reading tests have been developed to measure individual differences with respect to reading performance. The test items for these tests have usually been selected to maximally discriminate between individuals. That is, items with p values around .50 have been selected because they are most efficient for discriminating among individuals whereas items with p values approaching 1.00 or .00 have been eliminated because they do not have discriminating power. The research and development of the test also includes the establishment of a norm-group set of scores which is used as a standard to interpret individual differences on the test. Thus, the traditional measures of reading have been "norm-referenced."

The CDCRT differs from the traditional reading test in two ways. It was not developed to maximize individual differences, and test scores are not compared to a norm-group set of scores in the traditional fashion. Rather, the test was developed and revised so that: (a) a group of readers could not answer test items correctly before reading the passages, and (b) a group of readers could answer test items correctly after reading the passages. In operational terms, this meant that an item was revised if over 33% of a non-reading group answered it correctly or if less than 67% of a reading group answered it incorrectly. Thus, the test items were developed to discriminate between a reading group and a nonreading group, rather than among the readers of a reading group. The techniques of criterion-referenced measurement were used to develop the test so that when an individual gets an item correct, it is reasonable to infer that he stored this information while he was reading the passage. That is, a change has taken place as a result of reading.

The techniques of criterion-referenced measurement are also evident in the interpretation of the scores on the test. An individual receives three types of scores on the CDCRT. The most important score is the Efficiency score, i.e., the number of incorrect chunks, out of 100 possible, that the individual correctly identifies during the 25 minute test period. The other two scores are Rate and Accuracy. Rate is the number of the last item attempted and Accuracy is the percent correct, i.e., the Efficiency score divided by the Rate score. The CDCRT consists of five reading passages, each of which is immediately followed by a chunked test. Each test consists of 20, 5-alternative items where one alternative in each item is incorrect. Table 1 contains information extracted

from the section of the test manual concerned with interpretation of scores. As is evident in Table 1, each of the three test score variables has been dichotomized to produce the following categorizations of mature readers: Efficient-Inefficient, Rapid-Slow, and Accurate-Inaccurate. The Rapid-Slow cutoff point was selected to be 60-61. If an individual completed reading and answering not more than the first three passages and their corresponding test items (60 or less), then he was arbitrarily categorized as a slow reader. If an individual attempted Item 61, then he had also read Passage 4. Thus, he was probably more than a small increment faster than those individuals who had only finished 60 items. The Accurate-Inaccurate cutoff score was arbitrarily set at 67-68%, the same cutoff used for the initial revision of test items. The Efficiency-Inefficiency cutoff score of 40-41 was determined by multiplying the 67% accuracy by the 60 item rate. Thus, this cutoff allowed no overlap between the category of Efficient, Rapid, and Accurate readers and the category of Inefficient, Slow, and Inaccurate readers.

Insert Table 1 about here

Table 1 contains the six possible categories of mature readers. ("Mature reader" is a term used to describe the general type of reader who has progressed beyond the initial phases of learning to read and who is now reading to learn.) The Efficiency of a mature reader is made of two components, his Rate and Accuracy. Therefore a mature reader cannot be Rapid and Accurate and be classified Inefficient. At the same time, a reader cannot be Slow and Inaccurate and be Efficient. This would be a contradiction in terms. This situation is evidenced in Table 1 by the "Not Applicable" in cells referring to these inappropriate combinations.

The CDCRT test manual has not provided norm-group data with centiles for each possible score. Instead, the last column in Table 1 presents the percent of experimental subjects (N=143) who have fallen into each of the six categories of reader types (see Darby & Carver, 6). This is different from traditional norm-referenced measurement in that there is no suggestion that the group of experimental subjects is representative of any particular group in a normative sense. That is, although this group is a sample of mature readers, the sample was not systematically drawn with the intention that these percent values be representative of all mature readers or even all college students.

Traditional reading tests provide a continuous scale of individual differences with no suggestion of cutoff points to help diagnose inadequacies that might be successfully treated or to help researchers study certain types of readers. On the CDCRT, each individual has been placed into a category which suggests that his reading behavior may be like the other individuals in his category and different from the individuals in other categories. This categorization is tied to the absolute criteria of efficiency, rate and accuracy.

It should be recognized that there is no empirical research which supports the validity and fruitfulness of this particular categorization. However, the criterion-referenced aspects of the categorization suggest several interesting hypotheses that could be tested. Listed below are the six types of mature readers with several accompanying statements, which may be considered as speculative judgements or tenative hypotheses to be tested. It is obvious that research is needed to evaluate the following six types of mature readers.

Type I. Efficient, Accurate, Rapid Mature Readers are those readers who are most representative of the college population. That is, they tend to store and retrieve information at a rate and with a level of accuracy such that they probably have no difficulty with the normal course load in curricula requiring extensive amounts of reading. It is doubtful if this type of individual would benefit greatly from reading instruction.

Type II. Efficient, Accurate, Slow Mature Readers are probably the next most frequent type of college reader. That is, they are efficient in their reading, but their efficiency comes from their slow but sure performance rather than from their speed of reading. This type of reader is most likely to benefit from training designed to increase his reading rate. However, it remains to be demonstrated that the reading rate of a mature reader can be increased without a concommitant decrease in accuracy of the information stored.

Type III. Inefficient, Accurate, Slow Mature Readers are quite likely sacrificing speed for increased accuracy of retrieval. Yet, they may be reading at their own optimum or efficient rate. It may be difficult to improve their efficiency by reading instruction, although they are probably the most frequent type of reader who seeks such help. Their most desirable alternative may be to proceed through college with a load less than normal.

Type IV. Efficient, Inaccurate, Rapid Mature Readers probably sacrifice accuracy of information retrieval for an increase in amount of information exposure, although they still are efficient as arbitrarily defined by this test.

This individual may be able to improve or optimize his efficiency by decreasing his reading rate. Reading instruction for this individual may be wisely directed toward altering his reading strategy since he may be skimming material that requires reading.

Type V. Inefficient, Inaccurate, Rapid Mature Readers may be able to become efficient readers by learning to decrease their rate to improve accuracy. It seems likely that reading instruction for these individuals should put more emphasis upon making the reading rate fit the nature of the reading purpose. Accuracy and efficiency may automatically improve with a decrease in rate.

Type VI. Inefficient, Inaccurate, Slow Mature Readers are probably the least likely candidates to improve markedly with reading instruction. Their deficiency probably results from a lack of adequate background information or knowledge. If an individual is only of high school age, then he may be an inefficient reader simply because he is not a mature reader and further progression in school may provide the background necessary to become efficient, accurate, and rapid. If an individual is of college age or above, then the score may simply reflect that he has not stored the amount and type of verbal information that is necessary as a prerequisite. That is, he may not have the reading vocabulary or the background context necessary to store and retrieve the verbal concepts and ideas represented by the reading level of the test material. There are many possible causes for reading scores in this category, e.g., low motivation and inadequate vision. However, for many of these individuals, the most fruitful strategy for improving reading efficiency might be to gain more experience and background by reading more college level material.

In summary, the CDCRT has been developed through the application of the techniques of criterion-referenced measurement. This was accomplished by revising items on the basis of the performance of reading and nonreading criterion groups. Also, six types of hypothesized criterion groups have been developed to facilitate the interpretation of the individual scores on the test. Finally, the hypothesized attributes of the six types of mature readers have been presented in some detail in an effort to stimulate definitive research related to these hypotheses.

References

1. Carver, Ronald P. Analysis of "chunked" test items as measures of reading and listening comprehension. Journal of Educational Measurement, 1970, 7, 141-150.
2. Carver, Ronald P. What is reading comprehension and how should it be measured? In G. B. Schick (Ed.) Yearbook of the National Reading Conference, Milwaukee: National Reading Conference, 1970, in press.
3. Carver, Ronald P. Special problems in measuring change with psychometric devices. Proceedings of the A.I.R. seminar on evaluation research--strategies and methods. Pittsburgh: American Institutes for Research, 1970, 48-66.
4. Carver, Ronald P. & Darby, Charles A., Jr. Development and evaluation of a test of information storage during reading. Journal of Educational Measurement, 1970, in press.
5. Carver, Ronald P. & Darby, Charles A., Jr. Analysis of the Chunked Reading Test. Unpublished manuscript, 1970.
6. Darby, Charles A., Jr., & Carver, Ronald P. Manual for the Carver-Darby Chunked Reading Test. Washington: American Institutes for Research, 1970.
7. Popham, W. J. & Husek, T. R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6(1), 1-9.

Table 1

The Six Categories of Mature Readers

Accuracy	Rate	Efficiency	Six Types of Readers	Percent Readers
Accurate (68% or greater)	Rapid (61 or greater)	Efficient (41 or greater)	Type I	60%
		Inefficient (40 or less)	(Not applicable)	NA
	Slow (60 or less)	Efficient (41 or greater)	Type II	20%
		Inefficient (40 or less)	Type III	7%
Inaccurate (67% or less)	Rapid (61 or greater)	Efficient (41 or greater)	Type IV	3%
		Inefficient (40 or less)	Type V	3%
	Slow (60 or less)	Efficient (41 or greater)	(Not applicable)	NA
		Inefficient (40 or less)	Type VI	7%