

THE MEASUREMENT OF FACTORS OF COMPREHENSION IN READING

Ann M. Bezdek, University of Illinois

I. Introduction

Reading comprehension has long been considered to be one of the most important goals of reading instruction. Yet relatively little is known of the nature of this complex ability even though much research has been devoted to the study of the nature of reading comprehension in the past three decades.

It has been theorized that comprehension in reading is composed of many underlying skills or abilities and that in order to comprehend what is read an individual must utilize numerous varied mental processes. Although reading theorists have generally described the process of reading comprehension as consisting of a number of specific skills, the results of studies concerned with identifying component skills of comprehension in reading to date have proved to be inconclusive. Many of these studies have utilized factor analytic techniques and while some studies have concluded that reading comprehension can be regarded as a unitary skill, others have provided evidence to indicate that some comprehension skills may be specific and unique, thereby supporting a multi-factor theory of reading comprehension.

Interestingly, the majority of studies designed to investigate the components of comprehension in reading have examined comprehension skills in subjects at the high school or college level where the reading process is likely to be a complex and highly organized process. The importance of examining the nature of the process at the elementary school level has been emphasized by Lennon (1962), Davis (1968), Bezdek (1973) and others. The present study was designed to investigate the specific factors of comprehension in reading to determine the unique measurable aspects of reading comprehension at the elementary level. This paper describes the manner in which the research was conducted and the data analyzed. Section 2 is concerned with a review and synthesis of the more notable studies that employed a factor analytic solution to examine the component abilities of reading comprehension. Section 3 describes the skills selected for measurement and the methods used to develop the items. A description of the test administration procedures, and an explanation of the manner in which the items were analyzed and the test revised is the content of section 4. Section 5 is concerned with the experimental test results and an analysis of the subtests' ability to function as measures of separate abilities. The results of the factor analysis are also presented in this section. Finally, the conclusions are found in section 6.

II. Earlier Factor Analytic Studies of Reading Comprehension

One of the earliest factor analytic studies concerned with comprehension in reading was reported by Feder (1938). The Comprehension Maturity Test designed to measure the ability to answer factual questions, the ability to make inferences from the material read, appreciation of the material read, and speed of reading was administered to 99 college sophomores. The common variance of the variables was factor analyzed using Thurstone's centroid method of factor analysis. The results were then rotated orthogonally. Feder's analysis of the results led him to conclude that reading for information and reading for inference were two distinct reading abilities.

Two years after Feder reported his findings Gans (1940) designed a study to determine the relationship between general and critical reading comprehension. Several published tests were used as a measure of general reading comprehension, and as a measure of critical reading skills, Gans constructed the Test of Reading Selection-Rejection. The tests were administered to a group of 417 pupils in grades four, five, and six in two New York Schools. The centroid method of factor analysis was applied to the results and rotated to oblique solutions. Gans concluded that reference reading is composed of general reading ability, and the ability to determine the relevance of designated sentences, and that it also involved some function of delayed recall.

Langsam (1941) using eight published reading tests obtained the intercorrelations of 21 variables in a group of 100 female college freshman. After the application of the centroid method of factor analysis and a rotation of the axes, Langsam found four factors to be specific to the skill of reading comprehension. The factors were identified as a verbal factor concerned with word knowledge, a perceptual factor involved with the speed of the response, a word factor described as fluency in dealing with words, and a factor tentatively identified as seeing relationships. A fifth factor identified as a number factor was not basic to the reading process.

Perhaps the most notable of the factor analytic studies concerned with reading comprehension was reported by Davis (1941). Davis reviewed the literature and identified nine skills considered to be the most important by reading theorists:

1. Knowledge of word meanings.
2. Ability to select the appropriate meaning for a word or phrase in the light of its particular contextual setting.
3. Ability to follow the organization of a passage and to identify antecedents and references in it.
4. Ability to select the main thought of a passage.
5. Ability to answer questions that are specifically answered in a passage.
6. Ability to answer questions answered in a passage but not in words in which the question is asked.
7. Ability to draw inferences.
8. Ability to recognize the literary devices used in a passage and to determine its tone and mood.
9. Ability to determine a writer's purpose, intent, and point of view.

Davis constructed items to measure each specific skill and assembled the 240 items into a test which was administered to 421 freshmen college students in several eastern colleges. A principal-axis component analysis was performed on the matrix of obtained score variances and covariances of the nine skill scores. Davis found six skills to be independent: word knowledge, the ability to reason in reading, the ability to follow the organization of a passage, the ability to recognize literary devices, the ability to understand a writer's explicit statements, and the ability to identify the writer's purpose. Davis explained that only word knowledge and the ability to reason in reading were measured with sufficient reliability to warrant their use for practical purposes.

Conant (1942) investigated the nature of reading comprehension to determine if the skill is best represented as a unitary ability or if it is composed of numerous skills. She constructed a test to measure six skills and administered it to 256 high school students and 74 freshmen college students along with the Nelson Denny Reading Test and the American Council Psychological Exam. Conant found high intercorrelations of the subtest scores and after applying Hotelling's principal components method of factor analysis she maintained that reading comprehension is a unitary skill.

Hunt (1952) reexamined the factors involved in reading comprehension originally identified in Davis' 1941 study. A pool of items from the Cooperative Reading Tests were classified by a group of consultants according to the verbal definition of each skill. The resulting 224 item test was administered to 560 college students. After the tests were scored Hunt determined the internal consistency of each skill measure, the intercorrelations among the six skill measures, and finally applied a factor solution to appraise whether the variance common to the relationships among the skill measures could be accounted for by a general factor of reading comprehension. Hunt's results indicated

that reading comprehension is comprised of a general reading factor and a word factor.

In another experiment, Singer (1962) reported a study that explored the factors of reading ability at the fourth grade level. He constructed measures of 36 variables and administered the tests to 60 fourth grade pupils. Employing the "substrata-factor analysis" Singer identified three factors of reading; word recognition, work meaning, and reasoning in context.

More than 25 years after his original study, Davis (1968) reexamined the uniqueness of the comprehension skills he identified in 1941. Selecting items from a pool of items that were not used in the published form of the Cooperative Reading Comprehension Tests, Davis developed two forms of a test to assess comprehension skills. Following a trail testing, the items underwent a differential item analysis and from the original 40 items measuring each skill, 24 were selected for use in the final test. Final forms of the test were administered to 988 senior high school students. The test scores were used to perform a uniqueness analysis cross validated by items and by examinees. Davis found that two factors, vocabulary knowledge and drawing inferences, were relatively unique comprehension skills. These findings were similar to Davis' 1941 findings where two skills, memory for word meanings and reasoning in reading, provided the largest factor loadings.

Schriener (1968) sought to determine the extent to which various comprehension subtests actually measure unique aspects of reading comprehension at the elementary school level. Schriener either constructed items or selected items from published tests to measure eight skills. The experimental battery of tests was split into halves and both halves were administered to 513 fifth grade students. Four subtests from the Iowa Test of Basic Skills were included in the analysis. The intercorrelations between all 12 variables were computed and utilized to obtain an estimate of the uniqueness of the subtests. The reliabilities of the differences between all possible pairs of subtests were analyzed and the principal components method of factor analysis was applied and followed by a varimax rotation. Schriener found that speed of reading, listening comprehension, verbal reasoning, and speed of noting details were relatively independent skills. From the factor solution however, speed of reading was the only identifiable factor involved in reading comprehension.

From a review of the research concerned with identifying the components of reading comprehension it is apparent that there exists no conclusive evidence to support either a one factor or a multi-factor theory of reading comprehension. In several instances experimentors felt that evidence was secured to indicate a high degree of specificity from one skill

measure to another when in some cases skills were highly correlated and overlapped.

Despite the varying interpretations of the research findings it appears that there is some evidence of several factors in reading comprehension, a word factor, a reasoning factor, and a literal understanding factor have been indicated by several studies. The great disparity in the conclusions reached in these studies can be attributed to the various instruments employed to assess the skills in question, the various statistical techniques applied, the different combinations of skills selected for analysis, differences in characteristics and ages of the subjects, and the manner in which the empirical results were interpreted.

III. Skill Selection and Item Construction

After careful analysis of the empirical studies that have been concerned with the analysis of reading comprehension, six of the eight skills analyzed by Davis (1968) were selected for investigation:

- 1) Remembering word meanings
- 2) Inferring word meanings from the context
- 3) Understanding content stated explicitly
- 4) Understanding the main idea of a passage
- 5) Making inferences about the content
- 6) Recognizing the author's purpose and point of view

The first skill, remembering word meanings, involves vocabulary knowledge or the ability to understand word meanings. The second skill is concerned with the ability to determine the meaning of unknown words by using the context for clues. Understanding the content stated explicitly, skill three, is concerned with literal comprehension and involves the student's ability to comprehend facts and details. The fourth skill pertains to the ability to understand the main idea of a selection. Skills five and six can be termed critical reading skills: skill five requires the ability to interpret what is read and draw conclusions based upon the facts presented, while skill six involves the ability to determine why an author uses the words he does, to recognize his bias and prejudices, and to interpret his feelings.

It should be pointed out that Davis in 1968 examined these skills at the high school level and before investigating these skills at the intermediate grade level some proof of their appropriateness for examination at this level was needed. A review of studies concerned with reading and thinking abilities of elementary school children was undertaken and this review indicates that the six skills adopted for investigation were appropriate skills for analysis in the intermediate grades.

Items to measure each of the comprehension skills were carefully constructed to measure the skills under investigation. Each item was based on a separate passage at the appropriate level of difficulty in order to eliminate any spurious interrelationships of item scores.

After the items were constructed they were submitted to a group of reading and measurement experts for analysis and were subsequently revised or eliminated on the basis of the suggestions offered by the team of analysts.

The revised items were then assigned to either Form A or Form B, the two trial test forms of the Reading Comprehension Skills Test. Each trail form consisted of six subtests of 14 items each designed to measure the six reading comprehension skills.

IV. Trail Testing and Item Analysis

In March of 1972, each trail form of the Reading Comprehension Skills Test was administered to a group of 50 fifth grade students on consecutive days. Students participating in the tryout testing were located in four fifth grade classrooms in a public elementary school in Champaign, Illinois and were selected because of their availability. Students were told the purpose for reading before each subtest and were provided with several sample items. Ample time was provided for students to complete all items, but several students did not do so.

While taking the test students were monitored and the protocols of students who were blindly marking answers were excluded from the item analysis along with the protocols of students who failed to answer each item. After excluding these protocols, the scores from 45 students who took Form A and 47 students who took Form B were used for the purpose of item analysis.

To identify items for inclusion on the final experimental form of the Reading Comprehension Skills Test correlation coefficients between pass or fail on each item in Form A and the total scores for each subtest in Form A were computed. Each item was excluded from the total score of the subtest of which it was a part when it was correlated with that subtest score. In this manner, no coefficient was spuriously increased by including an item in the total score with which it was correlated. The same procedure was followed with the items in Form B, the other tryout test form.

The resulting correlations were carefully reviewed and from the 28 items measuring each of the six comprehension skills, 14 were selected for use on the final experimental test. The items selected for inclusion on the final test had a higher average correlation with the total score on the skill they were designed to measure than with the total scores on the other skills. These correlations are given in table 1. In

general the mean differences are small, as would be expected, but, nevertheless, the items for each skill had subjective and empirical justification for their inclusion.

TABLE 1
Median Correlation Coefficients of
Items with Total Subtest Scores

Skill	Median Correlation Coefficient of 14 items with total score of the skill measured by the items	Median Correlation Coefficient of 14 items with total scores of skills not measured by the items
1	.39	.29
2	.28	.27
3	.42	.33
4	.45	.37
5	.34	.31
6	.33	.31

To determine the appropriateness of the selected test items for fifth grade students, difficulty indices were calculated for each item. These indices are expressed as the proportion of students passing each item. The range of the difficulty indices and the median difficulty index for each skill test are shown in Table 2.

For efficiency in measurement items of appropriate difficulty are those with difficulty indices ranging from .30 to .70. Table 2 indicates that the median difficulty index for each skill measure falls within this range.

TABLE 2
Percentage of Students Marking Items
Correctly in Forms A and B

Skill	Range for 14 items	Median Percentage
1	331-81	63
2	36-79	67
3	51-83	65
4	43-69	63
5	40-78	61
6	36-72	49

In sum, item analysis data were used to identify the best items for inclusion in the final test form. The items selected were constructed and initially judged to measure the skill they were intended to measure. The items correlated more highly with the skill they were intended to measure than with the other skills and were of appropriate difficulty for fifth grade students.

V. Experimental Test Administration and Results

The final experimental test, Form C, of the Reading Comprehension Skills Test consisted of 84 items arranged into subtests of 14 items each to measure the six comprehension skills being

investigated. The sample for the final testing consisted of 369 fifth grade students enrolled in five public elementary schools in Champaign, Illinois. Previously administered standardized tests indicated that the students were slightly above average in general reading achievement. (Average reading achievement grade score was reported as 5.5 on the Scholastic Testing Service's Educational Development Test administered at the beginning of the school year.) The students comprising the sample for this study were selected mainly because of their availability and not because they were representative of fifth grade students throughout the state or country. Students participating in the final testing did not participate in the tryout testing.

Form C of the Reading Comprehension Skills Test was administered to the 369 fifth grade students in May of 1972. The test was administered in two one hour sessions on consecutive days and students were again provided with an explanation of the purpose for reading prior to taking each of the subtests. Sample items were also provided prior to subtest administration. Students were observed while taking the test and the protocols from a small number of students which were incomplete or incorrectly marked were excluded from the final analysis.

Following the scoring of the tests the first step in the analysis was to determine the reliability of the skill measures. The reliability estimates are based upon one administration of one test form, therefore, the reliability was determined through the use of a type of split-half method. The Kuder-Richardson formula 21 provided by Gronlund (1965 p.85) was used to obtain the coefficients. The resulting reliability coefficients are shown in Table 3 along with the standard error of measurement for each subtest and for the total test.

TABLE 3
Reliability* and Standard Error of Measurement
for the Six Skills on the Reading Comprehension
Skills Test Form C

Skill	r	S _m
1. Vocabulary	.56	1.77
2. Inferring Word Meaning	.64	1.72
3. Reading for Details	.80	1.47
4. Reading for the Main Idea	.75	1.61
5. Making Inferences	.72	1.64
6. Determining Author Purpose and Point of View	.70	1.72
Total Test	.94	4.05

*The estimate of reliability are based on the Kuder-Richardson Formula 21.

In general, the reliability coefficients for the six skill measures given in Table 3 are not sufficiently high to warrant any diagnostic interpretations of the test scores. It should be pointed out that the estimates in Table 3 are likely to be underestimates of the subtests' reliability since the KR-21 formula assumes that all items have equal difficulty. As a whole the test functions as a very accurate measure of reading comprehension.

The second step in the analysis was to obtain the intercorrelations among the skill measures. If each subtest in fact assesses a specific skill the intercorrelations among the various skill measures would tend to be low. Relatively high coefficients of correlation, on the other hand, would indicate that the subtests are measuring the same process. The intercorrelations among the subtests are given in Table 4.

TABLE 4
Intercorrelation Coefficients of Total
Scores on Subtests

	1	2	3	4	5	6
1						
2	.61					
3	.65	.70				
4	.66	.69	.75			
5	.64	.65	.70	.71		
6	.61	.61	.70	.70	.69	

The correlation matrix in Table 4 indicates coefficients above .60 and although what one calls "high" correlation is rather arbitrary, in this case it seems doubtful that the original hypothesis will be supported. To further investigate the uniqueness of the six skill measures the principal components method of factor analysis was applied to the matrix of intercorrelations given in Table 4. The results of the factor analysis are given in Table 5.

TABLE 5
Results of Factor Analysis Using Principal
Components Method

Skill	Factor I Test Loading	h^2	r	$r - h$
1	.811	.658	.560	-.098
2	.833	.694	.640	-.054
3	.885	.783	.800	.017
4	.882	.778	.750	-.018
5	.859	.738	.720	-.018
6	.842	.709	.700	-.009

An examination of Table 5 indicates that comprehension in reading for the group tested is a unitary skill. Although not shown in the table, the results of the factor analysis

indicates that nearly 73 percent of the total variance was accounted for by Factor I. The test loadings, or correlations of the skill scores with Factor I, are given in the second column of the table. It is apparent that all six skill measures are highly loaded on Factor I and due to the nature of the skill measures Factor I can tentatively be identified as general reading comprehension.

To examine the uniqueness of each skill measure in order to determine if there was any specific variance for each skill measure that was unaccounted for by the general factor, the reliability of each skill measure which indicates the maximum part of the total test variance that can be explained by that skill, given in column 4 of Table 5, was compared to its communality estimate, which reveals the percent of total variance accounted for by the general factor, given in column 3. The communality estimate when subtracted from the reliability estimate for each skill measure indicates the amount of specific variance unaccounted for. These values given in column 5 of Table 5, show that all of the test variance for each skill, given the reliability estimates, was accounted for by the general factor. Thus none of the skills have been found to be unique.

VI. Conclusion

The results of this study indicated that for the group tested, students who obtained high scores on one skill measure obtained high scores on the others as well. Thus there was a consistency of pupil ranking on the subtests and no evidence was obtained to indicate that students vary in their ability to utilize the six comprehension skills. Performance on the six skill measures was found to be a function of one general factor, tentatively identified as "general reading comprehension," which accounted for 73 percent of the total test variance. These results cannot be generalized to other groups of fifth grade students. It is possible that if the same items were administered to other groups of students a difference in the ability to apply each skill would be found. In this case the items would serve as differential measures of the six skills.

Suppose that subsequent administrations of the experimental test indicate that the subtests are not functioning to measure specific skills. Would this then imply that reading comprehension is a general unitary factor or that the component skills of reading comprehension are not the skills that have been investigated or that the measurement instrument was too crude? Until evidence is accumulated that leaves little doubt that reading comprehension is a general ability and is not composed of specific skills the search for the nature and specific skills involved in the process of comprehension in reading will undoubtedly continue.

The findings of this investigation did not provide evidence to support the multi-factor theory of reading comprehension. Because the analysis did not indicate any specific or unique test variance for the six skill measures no justification was provided to permit the interpretation of student scores on the subtests as indices of particular strength or weakness in the various skills. Thus the Reading Comprehension Skills Test could not be appropriately used as a diagnostic instrument, however, the findings suggest that the Reading Comprehension Skills Test is well suited to assess general reading comprehension ability. The total test reliability estimate indicates that the test accurately assesses comprehension ability and the items require students to read for a variety of purposes. It is desirable that general reading comprehension tests include items that require students to read for different purposes because of the influence testing has upon teaching. When items that require one to read for a variety of purposes appear on reading tests it is probable that instruction will be devoted toward developing these reading competencies. The Reading Comprehension Skills Test would be a valuable tool for these purposes.

The author knows of no other studies prior to the present one that investigated the components of reading comprehension at the elementary school level through the use of a test composed of items based upon separate passages that were carefully designed and constructed to measure specific comprehension skills. Future investigations of this type should continue to examine comprehension skills at the elementary school level and explore the construct validity of separate measures of comprehension skills included on experimental tests.

REFERENCES

- Bezdek, A. M. The Assessment of Fundamental Skills Involved in Reading Comprehension. Unpublished doctoral dissertation, University of Illinois, 1973.
- Conant, M. M. The Construction of a Diagnostic Reading Test, New York: Teachers College Bureau of Publications, Columbia University, 1942.
- Davis, F. B. Fundamental Factors of Comprehension in Reading. Unpublished doctoral dissertation, Harvard University, 1941.
- Davis, F. B. Fundamental Factors of Comprehension in Reading. Psychometrika, 1944, 9, 185-97.
- Davis, F. B. Research in Comprehension in Reading. Reading Research Quarterly, 1968, 3, 499-545.
- Feder, D. D. Comprehension Maturity Tests--A New Technique in Mental Measurement. Journal of Educational Psychology, 29, 597-606.
- Gronlund, N. E. Measurement and Evaluation in Teaching. Toronto: Macmillan Co., 1965.
- Hunt, L. C. A Further Study of Certain Factors Associated with Reading Comprehension. Unpublished doctoral dissertation, Syracuse University, 1952.
- Langsam, R. S. Factorial Analysis of Reading Ability. Journal of Experimental Education, 10, 57-63.
- Lennon, R. T. What Can be measured? The Reading Teacher, 1962, 15, 326-37.
- Schreiner, R. L. A Study of Interrelations Among Different Approaches to Measuring Reading Comprehension. Unpublished doctoral dissertation, University of Iowa, 1968.