

1983

## The Cloze Procedure: A Measure of Reading Comprehension Ability

Theresa Lorene Donaldson  
*University of North Florida*

Follow this and additional works at: <https://digitalcommons.unf.edu/etd>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Educational Methods Commons](#), [Elementary Education and Teaching Commons](#), [Higher Education Commons](#), and the [Secondary Education and Teaching Commons](#)

---

### Suggested Citation

Donaldson, Theresa Lorene, "The Cloze Procedure: A Measure of Reading Comprehension Ability" (1983). *UNF Graduate Theses and Dissertations*. 687.  
<https://digitalcommons.unf.edu/etd/687>

This Master's Thesis is brought to you for free and open access by the Student Scholarship at UNF Digital Commons. It has been accepted for inclusion in UNF Graduate Theses and Dissertations by an authorized administrator of UNF Digital Commons. For more information, please contact [Digital Projects](#).

© 1983 All Rights Reserved

THE CLOZE PROCEDURE:  
A MEASURE OF READING COMPREHENSION ABILITY

by


Theresa Lorene Donaldson


A thesis submitted to the Department of Elementary  
and Secondary Education in partial fulfillment  
of the requirements for the degree of  
Master of Education.

UNIVERSITY OF NORTH FLORIDA

COLLEGE OF EDUCATION

1983

  
Dr. Bruce Gutknecht, Advisor

  
Dr. James Mittelstadt, Committee

  
Dr. James Cangelosi, Committee

## Abstract

The present study was conducted in order to determine if a relationship exists between those comprehension processes measured by the Paragraph Comprehension subtest of the McGraw-Hill Basic Skills System Reading Test and the cloze test developed for this project. Two classes of undergraduate education majors enrolled in "Foundations of Performance-Based Education" were administered both the cloze test and the multiple choice comprehension test.

A correlation coefficient was computed between the summary scores on the McGraw-Hill and the summary scores on the cloze test. The results yielded a correlation of .021. These results did not support the research hypothesis.

To my family - for all the love  
and support they have given  
me throughout my years  
of study.

## TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION AND RATIONALE.....	1
	Problem Statement.....	2
	Definition of Terms.....	2
II.	A REVIEW OF RELATED LITERATURE.....	4
	Face Validity.....	4
	Concurrent Validity.....	5
	Construct Validity.....	10
III.	THE STUDY.....	16
IV.	RESULTS AND DISCUSSION.....	19
	Analysis of the Results.....	19
	Discussion.....	19
	Summary.....	20
V.	RECOMMENDATIONS FOR FURTHER STUDY.....	21
	REFERENCES.....	22
	APPENDIX - The Cloze Test.....	25

## INTRODUCTION AND RATIONALE

Measurement and evaluation play key roles in designing any developmental reading program. Without a valid method of assessing where a student's strengths and weaknesses lie, one cannot determine the goals of instruction or select suitable teaching strategies and appropriate instructional materials.

When a student takes a standardized reading comprehension test such as that found in Part III of the McGraw-Hill Basic Skills System Reading Test, his performance may be incorrectly evaluated for a number of reasons. Factors such as one's inability to deal with a multiple choice testing format, being required to apply a particular reading skill in an "artificial/abstract" setting, as well as operating within a limited time frame may prevent an accurate measurement of the student's true comprehension ability.

Since such factors may adversely affect the validity of the test data, it is desirable to look at another diagnostic tool to supplement the information gathered. The McGraw-Hill is currently being used to measure the reading comprehension ability of all entering education majors at an urban commuter university in northeast Florida. The use of an additional diagnostic tool would make more material available upon which students' strengths and weaknesses might be assessed.

The cloze test is being considered by reading specialists as an additional diagnostic tool for a variety of reasons. First, it provides a method for testing the reader's ability to use such reading strategies as context clues and to understand the syntactic position of words in sentences in order to overcome the obstacles to understanding the author's message. Furthermore, the cloze method tests these skills in the way in which they are used in the actual reading process. Another reason the cloze test should be considered a viable alternative to the conventional comprehension test is that it not only

measures a student's understanding of the message within a given text, but also his ability to integrate meaning across sentences so that he can understand the more extended messages found in text. This is accomplished because test items are passage dependent. That is, a student's ability to respond correctly to an item is totally dependent upon his understanding of the passage. An additional benefit to be gained from utilizing the cloze test as an evaluative tool is that some official testing programs include a cloze test. One such program is the Florida Teacher Certification Exam. Students in Florida teacher preparation programs should be exposed to this method in order to gain some proficiency in dealing with this particular testing format.

The relative merits of these two measurement tools--the standardized test as represented by the McGraw-Hill and the cloze test--have not been substantiated with certain groups of students. Therefore, the purpose of this research study was to determine if a relationship exists between those comprehension processes measured by the cloze test and the Paragraph Comprehension subtest of the McGraw-Hill Basic Skills System Reading Test when administered to the undergraduate education majors enrolled in "Foundations of Performance-Based Education" at the University of North Florida during the Summer Term of 1982.

#### Definition of Terms

Validity: the extent to which a test measures what it is supposed to measure or the accuracy of inferences drawn from test scores.

Content Validity: an inference about the test itself as a sample of an area of knowledge or skill.

Construct Validity: an inference about the test score as a measure of a construct of intellectual, perceptual, or psychomotor ability, motivation, or personality.

Concurrent Validity: an inference about an individual's standing on some other variable of interest available at the same point in time.

Phonology: the study of speech sounds that, when combined with other speech sounds, contribute to meaning and meaning variation.

Syntax: the set of principles by which words of a language are arranged in meaningful order.

Semantics: the study of word meaning involving the way in which users of a language divide reality into significant categories and label them in their language.



## A REVIEW OF RELATED LITERATURE

The purpose of this investigation was to determine if a relationship exists between those comprehension processes measured by the cloze test and the Paragraph Comprehension subtest of the McGraw-Hill Basic Skills System Reading Test. In order to obtain data on these problems, it is necessary to explore the validity of the cloze test as a measure of reading comprehension ability.

The cloze test will be explored in the review of the literature in terms of its face validity, its concurrent validity, and its construct validity. The face validity of the cloze test will be assessed in order to identify some similarities and differences between cloze tests and tests to which the label comprehension is conventionally applied. To determine if a relationship exists between those comprehension processes measured by cloze and conventional tests, the concurrent validity of the cloze test will be examined. The third section of this review will address the construct validity of the cloze test as a measure of reading comprehension ability in order to explore the nature of the relationship between cloze tests and conventionally made comprehension tests.

### Face Validity

Bormuth (1968) examined the similarities and differences between cloze tests and conventionally made comprehension tests and found that, while the cloze test may appear to be a completely different kind of test, there are some formal similarities between these measures of reading comprehension ability. He reported that many cloze test items are identical to conventional completion and multiple choice test items. He also pointed out that the processes involved in filling the cloze blanks are similar to those involved in answering conventionally made items. In both instances the subjects are presented with incomplete sentences and are required to fill the blanks by

guessing, from the context of remaining words, what the missing words should be.

Bormuth also noted that items generated by cloze and conventional test making procedures differ in several important ways. He stated these differences as follows:

1. In a cloze test, only one word is deleted at a time while in conventional tests, whole phrases and clauses may also be deleted. Furthermore, in cloze tests, structural words may be deleted while in conventional tests, only lexical words may be deleted by themselves.
2. Cloze tests are made only from sentences in the text while conventional test items may be made either from sentences in the text or from sentences that can derived from the text.
3. Cloze tests are taken by students who have not read the undeleted version while students taking conventional tests have been famaliarized with the printed material. (pp. 430-431) Note: This applies to cloze tests used for measuring the readability of a passage.

These findings, as reported by Bormuth, enable us to understand the basic similarities and differences between cloze and conventional reading comprehension tests. However, he suggests that perhaps too much emphasis has been placed upon the differences between these diagnostic tools. Bormuth points out that, because the student is given eighty percent of the text on which to base his answers, his ability to correctly fill the cloze blanks depends heavily upon an understanding of the text.

#### Concurrent Validity

Numerous studies have been conducted to determine if significant correlations exist between scores on cloze tests and tests to which the label comprehension is conventionally applied. Researchers have approached this question in a variety of ways.

Bornuth (1967) examined the literature on the relationship between cloze test scores and scores on conventional tests. He found a large number of studies reporting significant correlations between conventional tests and cloze tests made from the same passages. His findings are reported as follows:

Taylor, using Air Force trainees as subjects, found a correlation of .76;

Jenkinson, using high school students, found a correlation of .82;

Bornuth, using elementary school pupils, found correlations ranging from .73 to .84; and Friedman, who used college students, gave comprehension tests consisting of 8 to 12 items each and obtained correlations ranging from .24 to .43. (p. 8)

Bornuth suggests that these correlations are relatively high when taking into consideration the fact that, where test reliabilities were reported, the validity correlations and the reliabilities were approximately of the same proportion.

A number of researchers have explored the correlations between cloze tests and standardized tests of reading achievement. Jenkinson, cited in Rankin (1959), using the vocabulary and comprehension sections of the Cooperative Reading Test, found correlations of .78 and .73; Fletcher (1960) used the same test to assess the reading comprehension ability of college freshmen. The Cooperative Reading Test, Level of Comprehension showed a correlation of .55 while the Cooperative Reading Test, Vocabulary showed a correlation of .63.

Rankin (1958) used the Diagnostic Reading Test to measure the general reading ability of participants in a college reading improvement program. Correlations of .68 and .60 were reported between the cloze test scores and scores on the vocabulary and comprehension sections of the standardized test. Using the Stanford Achievement Test, Paragraph Meaning subtest, Ruddell (1963) found correlations which ranged from .61 to .74.

Gallant (1965) examined the use of the cloze test as a measure of readability of materials for primary grades. She obtained correlations ranging from .65 to .81 between cloze test scores and the Paragraph Reading section of the Metropolitan Achievement Test, Primary Batteries I and II.

Schneyer (1965) administered the vocabulary and comprehension sections of the Gates Reading Survey and a cloze test in which every tenth word was deleted. He reported a correlation of .74 with vocabulary and .68 with comprehension. When scores on a cloze test, using noun-verb deletions, were compared to scores on the standardized test, correlations of .63 and .60 were reported.

However, not all studies report significant correlations between cloze test scores and scores on conventional reading comprehension tests. A factor analysis study by Weaver and Kingston (1963) produced information contrary to these findings. Using the Davis Reading Test, they found correlations ranging from .25 to .51. The results indicated that the factors represented on both the comprehension and the cloze test had different loading patterns. The largest factor loadings occurred on separate factors with only slight to moderate loadings on the same factors. The correlations among various tests were submitted to a principal components analysis. The results indicated that the cloze test demonstrated low correlations with respect to the principal component. The comprehension test had its highest correlation with the same component. Based on these findings, Weaver and Kingston concluded that, as measures of reading comprehension ability, cloze tests do not appear to be very valid.

Bormuth (1969) disagreed with Weaver and Kingston's interpretation of this data by stating that it may be questioned for at least four reasons. First, because their subjects were college students, the range of variability was probably somewhat curtailed. Secondly, as Rankin pointed out in an earlier

study, their calculations were based upon correlations which differed in size from those obtained by other investigators. Third, unusual patterns of factor loadings were observed on the standardized test they used. Finally, some inconsistencies in the factor loading patterns among the cloze tests themselves were observed. Bormuth's analysis of the factor validity of cloze tests will be examined later in this review of the literature.

Several investigators have examined the inter-relationships between readability, cloze and multiple choice test scores. Significant correlations were reported between cloze and multiple choice test scores on passages whose readability values were determined by a number of different readability formulas.

Taylor (1953) examined the cloze procedure as a tool for measuring readability. The purpose of his study was to determine if cloze scores would rank three selected passages in the same order of readability as did the Flesch and Dale-Chall formulas. He found that the passages were consistently ranked in the same order of difficulty by the cloze readability procedure, the Flesch, and the Dale-Chall readability formulas.

Sukeyori, as reported by Bormuth (1968), compared the rankings of eight passages by both the cloze readabilities of passages and by the combined subjective rankings of the passages by three judges. He obtained a correlation of .83.

Bormuth (1967) sought to establish a frame of reference for interpreting the scores of cloze tests when they are used to measure the comprehension difficulties of passages. Fourth and fifth graders were tested on passages ranging in Dale-Chall readability from grades 4.5 to 6.5. The cloze test was given approximately three days before the multiple choice test. Bormuth reported a correlation of .95 between the two forms of the test. He found no evidence that scores on the multiple choice test were inflated because subjects had previously taken a cloze test over the same passage.

In a later study, Bormuth (1968) used four forms of the Gray Oral Reading Paragraphs. The results he obtained were similar to those reported in his previous study. When comparing the cloze readabilities of passages and the comprehension difficulties of the same passages, he found correlations ranging from .91 to .96. Correlations ranging from .90 to .95 were found between the cloze readabilities of passages and the word recognition difficulties of the same passages.

Rankin and Culhane (1969) explored the relationship between cloze and multiple choice comprehension test scores. Five passages ranging in Fry readability from 5.0 to 8.0 were administered to fifth grade students. Subjects took the cloze test first and the multiple choice test about four days later. They obtained correlations between the cloze and multiple choice comprehension tests ranging from .54 to .77.

In order to determine whether the frame of reference for interpreting cloze test scores established by previous researchers was applicable to high school and college age students, Peterson, Peters, and Paradis (1972) conducted a similar study. Cloze passages selected on the basis of a match between predictions of readability of a group of health education passages and grade equivalence scores on the Nelson-Denny Reading Comprehension Test were given to several groups of students. The results yielded cloze percentage levels which were equivalent to the 75 percent level on the multiple choice comprehension tests. The cloze percentage levels reported were 42 percent for high school students, 43 percent for adult-vocational technical students, and 44 percent for university students.

Another set of baseline data on cloze scores comes from Entin and Klare (1978). The purpose of their study was to compare scores on two forms of a cloze test, standard-sized solid blank and dashed-line blank, with multiple choice test scores and with readability formula values. Students were admin-

istered both a cloze test and a multiple choice test over seven passages taken from the Nelson-Denny Reading Test. The results of their study showed correlations between the cloze tests and the multiple choice test which were all highly significant. The dash cloze form yielded slightly higher percentages correct than the solid cloze form with correlations reported at .61 and .57 respectively.

When comparing the relative difficulty level of the seven passages as determined by the Flesch "Reading Ease" readability formula and by subjects scores on the multiple choice and cloze tests, the mean scores on both cloze forms correlated with readability values, but the mean scores on the multiple choice test correlated with neither. Therefore, a supplementary study was conducted to estimate how many multiple choice questions on the Nelson-Denny Reading Test could be answered correctly without reading the passages on which they were based. When the scores on the multiple choice test were corrected for this, they correlated positively with cloze test scores and with readability values.

The relationship between those comprehension processes measured by cloze tests and conventional comprehension tests has been examined in a variety of ways. Significant correlations have been reported between scores on: 1) cloze and conventional tests constructed from the same passages; 2) cloze and standardized tests of reading achievement; and 3) cloze and multiple choice test scores as compared with varying readability formula values. In general, these studies seem to show that a relationship does exist between the comprehension processes measured by these two diagnostic tools.

#### Construct Validity

The literature suggests that there is a sound theoretical basis for using cloze tests as measures of reading comprehension ability. Fries (1963) stated that language concerns itself with conveying meanings. He emphasized the fact

that the language itself does not constitute meaning but that it is a code of signals by which messages are transmitted from one individual to another.

Fries suggested that the English language is composed of several important layers of signaling patterns. He identified these layers as the signals of meaning carried by lexical items, the signals of meaning carried by grammatical structures, and the signals of social-cultural meanings.

The cloze procedure measures a reader's ability to use all three layers of language meaning. When speaking in terms of successfully reconstructing a message, McLeod (1965) stated that the process

requires a familiarity with the grammatical structure of English, an understanding of lexical meaning and, if the passages selected are concerned with a variety of experiences in a given culture, they reflect to some extent social-cultural meaning. (p. 3)

Davis (1968) provided further support to the idea that comprehension involves an interaction between all three layers of language meaning. He argued that comprehension is not a unitary construct.

The miscue research has indicated that language is composed of three interrelated systems--or cue systems--which provide the reader with strategies enabling him to overcome the obstacles of understanding the author's message. When speaking in terms of oral language, these cue systems are referred to as the graphophonic, the syntactic, and the semantic systems.

Menosky (1972), when talking about the graphophonic cue system, stated that reading entails more than matching letters and sounds; it involves the complex psycholinguistic process of reading. She further added that a particular cue system's importance is determined by its relation to and interaction with the other cue systems.

This view is clearly shared by Page (1971). He suggested that, in order to successfully reconstruct the author's message, a reader must be able to



interrelate all three language systems. This basic understanding, according to Page, should be reflected in our teaching of reading. Potential readers must be taught to respond to the three types of information interrelatedly and simultaneously in their search for meaning.

Burke (1970) described the syntactic system as a kind of exchange through which all three language systems interact. It is added that, while the syntactic system may be somewhat unique to language, the three language systems function fully only when they function interdependently within the reading process.

This idea is extended to her thoughts on the evaluation process. When addressing the current methods of assessing reading comprehension ability, Burke believed that many reading tests hinder the function of one or more of the language systems by establishing minimal reading situations. She explained that language systems must be kept within the context of the ongoing reading process in order to be accurately measured and examined.

In view of these findings, the question arises: "When assessing a student's reading comprehension ability, is reading being measured as an inter-related process or is the reader being required to apply various reading strategies out of the context in which he/she would normally apply them?" Researchers have asked similar questions and have sought to determine what is actually being measured by reading comprehension tests.

A study conducted by Kendal, Mason, and Hunter (1980) examined the interaction between passage and task in order to determine which, if any, of the operational definitions used in test construction is actually measuring reading comprehension ability. Differences in fifth graders' reading comprehension scores were obtained using four different tasks generally used to measure comprehension and four different reading passages that were equated to readability formulas. The conclusion reached was that each of the tests

measures comprehension, and yet, at the same time, none measures it as an interrelated process. Their results suggested serious limitations of most reading comprehension research and testing.

Bormuth (1969) explored the factor validity of the cloze test. Nine passages of approximately 250 words each were used in the study. They were distributed in Dale-Chall readability from 4.5 to 6.5 and were drawn from a number of subject matter areas. A 50 item cloze test was made from each passage by replacing every fifth word with an underlined blank of standard length. Seven multiple-choice tests were also constructed from these passages. The tests were written to measure comprehension of vocabulary, of explicitly stated facts, of sequence of events, of stated casual relationships, of the main idea of passages, of inferences, and of the author's purpose. The tests were administered to 150 students in grades four, five, and six who constituted random samples of the populations of children at those grade levels in a single school district.

The correlations between the cloze and conventional tests over each passage ranged from .73 to .84. When the correlations were corrected for the unreliabilities of the tests, the correlations approached 1.00. These data were interpreted as providing little grounds for claiming that cloze tests measure anything other than what has commonly been referred to as reading comprehension skills. In addition, it may be concluded that cloze tests measure a variety of comprehension responses.

In an attempt to further explore the complex issue of testing to determine reading levels of students, Davidson (1980) conducted a study which compared the results from an informal reading inventory (IRI), a standardized test (SRA), a group reading inventory (GRI), and a cloze test. Comparisons were made using test scores and reading levels. Thirteen fourth grade classes and eleven sixth grade classes from schools in Northeast Arkansas and

Southeast Missouri were used in the study. A total population of 624 students were administered a cloze test in science and one in social studies, a group reading inventory in science and one in social studies, and a standardized reading test. Informal reading inventories were administered to 25 randomly selected students in three fourth grade classes, and to 25 randomly selected students in three sixth grade classes, for a total of 50 students.

When analyzing test scores for grades four and six, the cloze and GRI tests were found to be more highly correlated with one another than with SRA. The SRA correlated markedly stronger with the cloze and GRI tests than did the IRI test scores. Factor analysis showed that for both grades four and six, cloze science, cloze social studies, GRI science, and GRI social studies loaded on Factor 1, which was labeled "content reading and concept processing". These tests were based on general reading ability and extended to the ability to process specialized vocabulary and knowledge of the content areas of science and social studies. The IRI and SRA loaded heavily on Factor 2, which was labeled "general reading". These tests were based on general educational factors for reading plus verbal ability.

Based on these findings, the following conclusions were reached:

- 1) The diagnosis of reading ability needs to go beyond the measurement of general reading ability if the user intends to examine the reader's ability to apply the process of reading to content materials.
- 2) Students may be performing poorly in content subjects not because they lack general reading ability, but because they lack the ability to apply the necessary reading skills for content reading.

A number of researchers (Davis, 1941; Humber, 1942; Shores, 1960; and Swenson, 1961), cited in Maney (1958), have produced experimental evidence at the secondary school level which suggests that comprehension of reading materials is different in every subject area. They found that: 1) the skills

to be applied in a particular reading situation are determined by the content with which one is dealing and the purpose for reading; 2) the success one experiences when reading in one content area does not necessarily guarantee success in another; and 3) the ability to interpret content literally does not insure equal ability when working with higher levels of interpretation. They too concluded that the diagnosis of a student's reading performance in a content area must extend beyond the general reading ability. This position was further supported by Farr (1969) who stated that:

The diagnosis of reading ability needs to go beyond an evaluation of general reading power and should examine the reader's ability to apply his reading skills. (p. 117)

The cloze test has been examined in this portion of the literature review in terms of its construct validity. Research has provided a sound theoretical basis for its use as a measure of reading comprehension ability.

In summary, the literature indicates that the cloze test extends beyond the measurement of general reading skills. It requires the reading of content material plus the processing of concepts and principles dealing with content materials. Furthermore, the cloze test measures a reader's ability to apply such reading strategies as context clues and to understand the syntactic position of words in sentences within the context of the ongoing reading process. Based on these findings, it may be concluded that the cloze test could provide us with a relevant method of assessing reading comprehension ability.

## THE STUDY

### The Sample

The study was conducted at an upper-division, urban, commuter university in northeast Florida. The average age of those students attending this institution was approximately 30. The sample for this investigation was composed of two classes of students enrolled in "Foundations of Performance-Based Education". There were 13 students tested in the morning class and 25 students tested in the evening class which yielded a total of 38 students who participated in the study. Since the sample of students was comprised of members of intact classrooms, it seemed reasonable to assume that a representative sample had been taken and hence, the results of the study could be generalized to most students enrolled in the course.

### Measurement Instruments

The following instruments were used in the study:

McGraw-Hill Basic Skills System Reading Test (Paragraph Comprehension subtest; 1970)

The Paragraph Comprehension subtest of the McGraw-Hill contains five long reading passages. Comprehension of each of these passages is assessed by five multiple choice items, each containing four alternatives. These items measure the following skills: recognition of the main idea, recognition of details, understanding of general principles, discovery of paragraph organization and structure, and evaluation of tone and intent. Interspersed among the five longer passages are five shorter paragraphs with one item each. These items are designated to measure factual knowledge. The difficulty level of these passages as a whole centers on Grades 10 through 13, while the content comes mainly from college academic disciplines.

The tests were administered and scored according to the directions in the manual. In answering the questions, the subjects were allowed to refer

to the passages. A time limit of 40 minutes was set for test completion.

### Cloze Procedure

The cloze procedure is a test for measuring the effectiveness of communication from the writer to the reader. The term "cloze" is derived from the Gestalt concept of "closure," which is the attempt to complete a structural whole by filling in what is missing. The cloze procedure requires the systematic deletion of every "nth" word in a prose selection and evaluates how accurately the reader supplies the words that were deleted (Bormuth, 1968).

The five passages which comprised the cloze test used in this study were taken from Reading Power and Study Skills for College Work (Lefevre and Lefevre, 1978). The passages dealt with such subjects as American History, Education, Psychology, and Political Science. A complete set of statistics on the readability of the passages as provided by the Estimated Dale, the Fog Index, the Flesch Grade Level, and the Smog Index are shown in the following table:

TABLE I

	<u>Passage I</u>	<u>Passage II</u>	<u>Passage III</u>	<u>Passage IV</u>	<u>Passage V</u>
<u>Dale</u>	6.1	6.0	5.5	5.4	6.8
<u>Fog</u>	14.3	16.0	11.6	9.7	15.7
<u>Flesch</u>	10.0	17.5	9.3	7.6	13.9
<u>Smog</u>	9.8	10.9	9.4	8.4	9.3
<u>Average</u>	10.1	13.2	8.8	7.6	12.2

As indicated above, the difficulty level of the passages ranged from approximately Grades 8 through 13. Each passage was approximately 150 words in length, with each containing about 25 blanks. The first and last sentences

of the selections were left intact. In the other sentences, every seventh word was deleted and was replaced with an underlined blank of standard length (15 type spaces).

Each test booklet was comprised of five complete passages. The answer booklets contained the same five passages in cloze form. Students were instructed to: 1) choose three of the five passages which were of the most interest to them; 2) read each selection carefully; and 3) fill the deleted version of the passage without referring to the completed version. A suggested time of 45 minutes was given for the completion of this task. When scoring reader responses on the cloze test, exact word replacements were considered correct. Synonyms for these words and unambiguous misspellings were also counted correct.

#### Procedures

Experimental subjects were administered both the cloze test and the multiple choice test. The tests were given exactly one week apart, with the multiple choice test being the first one given in the morning class followed by the cloze test. The procedure was reversed for the evening class. Both cloze tests were based on the same five passages. As previously noted, a generous time of 45 minutes was allowed to complete the cloze test and a time of 40 minutes was allowed for the completion of the multiple choice test.

## RESULTS AND DISCUSSION

### Analysis of the Results

For the purpose of comparing the results from the administration of these diagnostic tools--the McGraw-Hill Basic Skills System Reading Test and the cloze test developed for this project--those subjects who did not take both tests were eliminated from the study. These were subjects who took the McGraw-Hill, which was a course requirement, but did not choose to take the cloze test. Of the original 44 subjects enrolled in the two sections of "Foundations of Performance-Based Education," six subjects were thereby eliminated.

To determine whether the testing procedure used in this study had any inflationary effect on subjects' scores, the summary scores of the experimental subjects, who took the multiple choice test first and then the cloze test, were compared to the summary scores of the experimental subjects who took the cloze test and then the multiple choice test. A comparison of the mean scores for the two groups showed that both groups of subjects performed significantly better on the cloze test than on the McGraw-Hill, regardless of the order in which they took the test. The mean scores were reported at .91 and .74 respectively.

A correlation coefficient was computed between the summary scores on the McGraw-Hill and the summary scores on the cloze test. The results yielded a correlation of .021. Because a significant correlation did not exist, further data analyses were not performed.

### Discussion

These results were not in agreement with many of the previous research studies on this subject. However, several important factors may have contributed to the results found in this study. Perhaps the most important factor was that, while both of the diagnostic tools used in this study were



designed to measure reading comprehension ability, each measures it quite differently. As Davidson (1980) pointed out, the cloze test measures content reading and concept processing, whereas, the multiple choice test measures general reading ability. The implication is that the cloze test measures one type of reading skill, while the multiple choice test measures a different type of reading skill. Another factor which may have influenced this study's results was the selection of topics. Perhaps a wider range of topics would enhance our ability to obtain an accurate measure of the students' reading comprehension ability. Finally, in response to the findings of Weaver and Kingston (1963), Bormuth (1969) stated that the low correlations between cloze and multiple choice test scores may have resulted from the fact that their subjects were college students, therefore, the range of variability was probably somewhat curtailed.

#### Summary

The present study was conducted to determine if a relationship exists between those comprehension processes measured by the cloze test and the Paragraph Comprehension subtest of the McGraw-Hill Basic Skills System Reading Test when administered to two classes of undergraduate education majors enrolled in "Foundations of Performance-Based Education." A correlation coefficient was computed between the summary scores on the McGraw-Hill and the summary scores on the cloze test. The results yielded a correlation of .021. These results did not support the research hypothesis.

RECOMMENDATIONS FOR FURTHER STUDY

The present study examined the cloze procedure as a technique for measuring reading comprehension ability. The literature indicates that the cloze test and the multiple choice comprehension test provide us with different measures of comprehension ability. Therefore, it may be recommended that both measures be used to assess the reading comprehension ability of the undergraduate education majors enrolled in "Foundations of Performance-Based Education." This would make additional material available upon which students' strengths and weaknesses might be assessed. In addition to this recommendation, it is suggested that a follow-up study be conducted to determine which, if any, measure of reading comprehension ability is a more accurate predictor of G. P. A. in the undergraduate education program.



REFERENCES

- Bormuth, J. R. Implications and use of the cloze procedure in the evaluation of instructional programs. Occasional Report No. 3. Center for the Study of Evaluation Instructional Programs, University of California, Los Angeles, 1967, p. 8.
- Bormuth, J. R. Comparable cloze and multiple-choice comprehension test scores. Journal of Reading, 1967, 10, 291-299.
- Bormuth, J. R. Cloze test readability: criterion scores. Journal of Educational Measurement, 1968, 5, 189-196.
- Bormuth, J. R. The cloze readability procedure. Elementary English, 1968, 45, 429-436.
- Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension ability. Reading Research Quarterly, 1969, 4, 358-365.
- Burke, C. L. A psycholinguistic description of grammatical restructurings in the oral reading of a selected group of middle school children. Doctoral dissertation, Wayne State University, Ann Arbor, Mich.: University Microfilms, 1970. (No. 70-03416).
- Davidson, E. S. An analysis of fourth and sixth grade reader performance using cloze tests, group reading inventories, an informal reading inventory, and a standardized reading test. ERIC Research in Education, Nov. 1980. (ED 195-984).
- Davis, F. B. Research in comprehension in reading. Reading Research Quarterly, 1968, 4, 499-545.
- Entin, E. B. & Klare, G. R. Some inter-relationships of readability, cloze and multiple choice scores on a reading comprehension test. Journal of Reading Behavior, 1978, 10(4), 417-436.
- Farr, R. C. Reading: what can be measured? Newark, Delaware: International Reading Association Research Foundation, 1969.

- Fletcher, J. E. A study of the relationships between ability to use context as an aid in reading and other verbal abilities. Dissertation Abstracts, 1960, 20, (No. 2675-6A).
- Fries, C. C. Linguistics and reading. New York: Holt, Rinehart, and Winston, 1963.
- Gallant, R. An investigation of the use of the cloze test as a measure of readability of materials for the primary grades. Dissertation Abstracts, 1965, 25, (No. 6431A).
- Kendal, J. R., Mason, J. M. & Hunter, W. Which comprehension? Artifacts in the measurement of reading comprehension. Journal of Educational Research, 1980, 73(4), 233-235.
- Lefevre, C. A. & Lefevre, H. E. Reading power and study skills for college work.
- Maney, E. S. Literal and critical reading in science. Journal of Experimental Education, 1958, 27, 57-64.
- McLeod, J. Gap reading comprehension test manual. Melbourne: Heinemann, 1965.
- Menosky, D. M. A psycholinguistic analysis of oral reading miscues generated during the reading of varying portions of text by selected readers from grades two, four, six, and eight. Doctoral dissertation, Wayne State University, Ann Arbor, Mich.: University Microfilms, 1972. (No. 72-14600).
- Page, W. D. A psycholinguistic description of patterns of miscues generated by a proficient reader in second grade, an average reader in fourth grade, and an average reader in sixth grade encountering ten basal reader selections ranging from preprimer to sixth grade. Doctoral dissertation, Wayne State University, Ann Arbor, Mich.: University Microfilms, 1971. (No. 71-17298).
- Peterson, J., Peters, N., & Paradis, E. Validation of the cloze procedure as a measure of readability with high school, trade school, and college pop-

- ulations. In F. B. Green (Ed.), Investigations relating to mature reading. Twenty-first Yearbook of the National Reading Conference. Milwaukee, Wisc.: National Reading Conference, 1972, 45-50.
- Rankin, E. F. An evaluation of the cloze procedure as a technique for measuring reading comprehension. Dissertation Abstracts, 1958, 19, (No. 733-4A).
- Rankin, E. F. The cloze procedure- its validity and utility. In Causey, O. S. and Eller W. (Eds.). Eighth Yearbook of the National Reading Conference, I. R. C., Milwaukee. 1959, 138.
- Rankin, E. & Culhane, J. W. Comparable cloze and multiple choice comprehension test scores. Journal of Reading, 1969, 13, 193-198.
- Ruddell, R. B. An investigation of the effect of the similarity of oral and written patterns of language structure on reading comprehension. Unpublished doctoral dissertation, Indiana University, 1963.
- Schneyer, J. W. Use of the cloze procedure for improving reading comprehension. The Reading Teacher, 1965, 19, 174-179.
- Taylor, W. L. Cloze procedure: a new tool for measuring readability, Journalism Quarterly, 1953, 30, 415-433.
- Weaver, W. W. & Kingston, A. J. A factor analysis of the cloze procedure and other measures of reading and language ability, Journal of Communication, 1963, 13, 252-261.

APPENDIX

The Cloze Test

From: Reading Power And Study Skills For College Work  
by Carl A. Lefevre and Helen E. Lefevre

Paragraph 1

The Formation of Colonies

Vermont was settled entirely from the other states of New England. The inhabitants have, of course, the New England character, with no other difference beside what is accidental. In the formation of colonies, those who are first inclined to emigrate are usually such as have met with difficulties at home. These are commonly joined by persons, who, having large families and small farms, are induced for the sake of settling their children comfortably to seek for new and cheaper lands. To both are always added the discontented, the enterprising, the ambitious, and the covetous. Many of the first and some of all these classes are found in every new American colony during the first ten years after its settlement has commenced. From this period on, kindred, friendship, and former neighbors prompt others to follow them. Still others are allured by the prospect of gain, presented to the sagacious in every new country from the purchase and sale of new lands; while not a small number are influenced by the brilliant stories which everywhere are told concerning most tracts during the early progress of their settlement.

Paragraph 2

An Urban Community College

An urban community college generally offers its students a variety of two-year vocational curricula, ranging from nursing to photography to secretarial science. Each two-year curriculum requires approximately sixty hours of credit, or four full semesters of college work. These two-year programs lead to an A.A. degree or certificate from the college. Many students find



the vocational programs helpful in broadening their education, since all require some courses such as English and sociology; moreover, such programs are designed to lead into specific job opportunities. Other students attend community colleges to enter transfer programs, earning credits that may be transferred to baccalaureate institutions for application toward a B.A. degree. A student who has selected a transfer institution should discuss his plans with his curriculum adviser or counselor. It may also be advisable to visit the transfer institution, study its catalogue, and consult its admissions office. The acceptance of transfer credits is the prerogative of the receiving institution. Because the urban community college offers both two-year vocational curricula and transfer programs, it attracts a variety of students.

Paragraph 3

Stress

We are all under stress, and from a fairly early age. Children are afraid of the dark, of unfriendly dogs, of losing their parents' love, of the neighborhood bully; and many adults have similar fears. In addition to fear, another kind of stress is anxiety. Students are anxious about their grades and their dates and their future plans; parents feel anxiety about their children's health and success in life. Both the young folks and the adults may feel another kind of anxiety, a feeling of inadequacy, perhaps about their intelligence or their ability to handle their life situations. Many everyday occurrences cause passing stress - events like driving in bad traffic, a sudden storm, a quarrel, a frown from the boss, a feeling of guilt or anger. More serious matters, such as a death in the family, a divorce, a serious health problem, loss of hard-earned property, cause deeper and longer-lasting stress. Other serious events, such as marriage, getting

a great new job, sudden wealth, a joyous reunion with a loved one, can be stressful too. Yes, we are all under stress. Yet most of us manage to keep our cool to quite an extent, and keep going in spite of all the stresses of our daily lives.

#### Paragraph 4

##### The Montessori Method

Children who enter a Montessori class for the first time begin to work on a wide range of activities related to real life. They discover how to manipulate shoelaces, buckles, snaps and bows by practicing with these objects mounted on small wooden frames. They learn to serve juice, scrub their hands, clean their work area when they are finished, and move their chairs when sitting or rising. These jobs are not intended solely to teach a youngster domestic chores. "Children experience joy at each fresh discovery," said Dr. Montessori. "Their satisfaction encourages them to seek new sensations and discoveries." Preparation for such tasks is in the spirit of Dr. Montessori's edict: "Teach the importance of doing even the smallest tasks well." Through expanding abilities gained in these early assignments, children begin to see order in apparent confusion. They begin to learn how to acquire the independence that comes from working for oneself. They begin to learn how to start and finish a job. Perhaps most important, they begin to understand what they can do.

#### Paragraph 5

##### Thomas Paine and the American Public

Tom Paine's pamphlet Common Sense, written in 1776, was a great call to arms; it was immensely influential in preparing the American public for the fight for independence from England. Later his Crisis papers were read a-

round the campfires of the revolutionary army, on the orders of Washington, to keep up the spirits of the men in the darkest days of the war. Paine was the pamphleteer of the revolution, praised by the leaders of the new nation, beloved by his countrymen. But The Age of Reason, written in France, with its disapproval of church or institutionalized religion, and its praise of "natural religion," turned the American public against him. Actually, like Franklin, Paine was a religious man who happened to believe in natural religion, not in the conventional religion of church or synagogue. Yet years after his death, he was still regarded with honor. Natural religion was considered to be the same as atheism. This was hardly logical, since the clearly stated purpose of his book was to fight against atheism, which was growing in France. He wanted to show that there was a real and valuable natural religion available to replace the older religion of the Bible, which many Frenchmen found they could no longer believe in. He repeatedly stated his belief in "one god" the creator. When Paine expressed his belief in independence, going with the trend of the time, he was praised by the public; when he expressed his belief in natural religion, going against common beliefs, he was damned.

From: Reading Power And Study Skills For College Work  
by Carl A. Lefevre and Helen E. Lefevre

Paragraph 1

The Formation of Colonies

Vermont was settled entirely from the other states of New England. The inhabitants have, of course, the \_\_\_\_\_ character, with no other difference beside \_\_\_\_\_ is accidental. In the formation of \_\_\_\_\_, those who are first inclined to \_\_\_\_\_ are usually such as have met \_\_\_\_\_ difficulties at home. These are commonly \_\_\_\_\_ by persons, who, having large families \_\_\_\_\_ small farms, are induced for the \_\_\_\_\_ of settling their children comfortably to \_\_\_\_\_ for new and cheaper lands. To \_\_\_\_\_ are always added the discontented, the \_\_\_\_\_, the ambitious, and the covetous. Many \_\_\_\_\_ the first and some of all \_\_\_\_\_ classes are found in every new \_\_\_\_\_ colony during the first ten years \_\_\_\_\_ its settlement has commenced. From this \_\_\_\_\_ on, kindred, friendship, and former neighbors \_\_\_\_\_ others to follow them. Still others are allured by the prospect of gain, presented to the sagacious in every new country from the purchase and sale of lands; while not a small number are influenced by the brilliant stories which everywhere are told concerning most tracts during the early progress of their settlement.

Paragraph 2

An Urban Community College

An urban community college generally offers its students a variety of two-year vocational curricula, ranging from nursing to photography to secretarial science. Each two-year curriculum requires approximately sixty \_\_\_\_\_ of credit, or four full semesters \_\_\_\_\_ college work.

These two-year programs lead \_\_\_\_\_ an A.A. degree or certificate from \_\_\_\_\_ college. Many students find the vocational \_\_\_\_\_ helpful in broadening their education, since \_\_\_\_\_ require some courses such as English \_\_\_\_\_ sociology; moreover, such programs are designed \_\_\_\_\_ lead into specific job opportunities. Other \_\_\_\_\_ attend community colleges to enter transfer \_\_\_\_\_, earning credits that may be transferred \_\_\_\_\_ baccalaureate institutions for application toward a \_\_\_\_\_ degree. A student who has selected \_\_\_\_\_ transfer institution should discuss his plans \_\_\_\_\_ his curriculum adviser or counselor. It \_\_\_\_\_ also be advisable to visit the \_\_\_\_\_ institution, study its catalogue, and consult \_\_\_\_\_ admissions office. The acceptance of transfer \_\_\_\_\_ is the prerogative of the receiving \_\_\_\_\_. Because the urban community college offers both two-year vocational curricula and transfer programs, it attracts a variety of students.

Paragraph 3

Stress

We are all under stress, and from a fairly early age. Children are afraid of the dark, \_\_\_\_\_ unfriendly dogs, of losing their parents' \_\_\_\_\_, of the neighborhood bully; and many \_\_\_\_\_ have similar fears. In addition to \_\_\_\_\_, another kind of stress is anxiety. \_\_\_\_\_ are anxious about their grades and \_\_\_\_\_ dates and their future plans; parents \_\_\_\_\_ anxiety about their children's health and \_\_\_\_\_ in life. Both the young folks \_\_\_\_\_ the adults may feel another kind \_\_\_\_\_ anxiety, a feeling of inadequacy, perhaps \_\_\_\_\_ their intelligence or their ability to \_\_\_\_\_ their life situations. Many everyday occurrences \_\_\_\_\_

\_\_\_\_\_ passing stress - events like driving in \_\_\_\_\_ traffic, a sudden storm, a quarrel, \_\_\_\_\_ frown from the boss, a feeling \_\_\_\_\_ anger or guilt. More serious matters, \_\_\_\_\_ as a death in the family, \_\_\_\_\_ divorce, a serious health problem, loss \_\_\_\_\_ hard-earned property, cause deeper and longer-lasting \_\_\_\_\_. Other serious events, such as marriage, \_\_\_\_\_ a great new job, sudden wealth, \_\_\_\_\_ joyous reunion with a loved one, \_\_\_\_\_ be stressful too. Yes, we are \_\_\_\_\_ under stress. Yet most of us manage to keep our cool to quite an extent, and keep going in spite of the stresses of our daily lives.

#### Paragraph 4

#### The Montessori Method

Children who enter a Montessori class for the first time begin to work on a wide range of activities related to real life. They discover how to manipulate shoelaces, \_\_\_\_\_, snaps and bows by practicing with \_\_\_\_\_ objects mounted on small wooden frames. \_\_\_\_\_ learn to serve juice, scrub their \_\_\_\_\_, clean their work area when they \_\_\_\_\_ finished, and move their chairs quietly \_\_\_\_\_ sitting or rising. These jobs are \_\_\_\_\_ intended solely to teach a youngster \_\_\_\_\_ chores. "Children experience joy at each \_\_\_\_\_ discovery," said Dr. Montessori. "Their satisfaction encourages \_\_\_\_\_ to seek new sensations and discoveries." \_\_\_\_\_ for such tasks is in the \_\_\_\_\_ of Dr. Montessori's edict: "Teach the importance \_\_\_\_\_ doing even the smallest task well." \_\_\_\_\_ expanding abilities gained in these early \_\_\_\_\_, children begin to see order in \_\_\_\_\_ confusion. They begin to acquire the \_\_\_\_\_ that comes with working for oneself. \_\_\_\_\_ begin to learn

how to start \_\_\_\_\_ finish a job. Perhaps most important, they begin to understand what they can do.

### Paragraph 5

#### Thomas Paine and the American Public

Tom Paine's pamphlet Common Sense, written in 1776, was a great call to arms; it was immensely influential in preparing the American public for the fight for independence from England. Later his Crisis papers were read \_\_\_\_\_ the campfires of the revolutionary army, \_\_\_\_\_ the orders of Washington, to keep \_\_\_\_\_ the spirits of the men in \_\_\_\_\_ darkest days of the war. Paine \_\_\_\_\_ the pamphleteer of the revolution, praised \_\_\_\_\_ the leaders of the new nation, \_\_\_\_\_ by his countrymen. But The Age of Reason, written \_\_\_\_\_ France, with its disapproval of church \_\_\_\_\_ institutionalized religion, and its praise of " \_\_\_\_\_ religion," turned the American public against \_\_\_\_\_. Actually, like Franklin, Paine was a \_\_\_\_\_ man who happened to believe in \_\_\_\_\_ religion, not in the conventional religion \_\_\_\_\_ church or synagogue. Yet years after \_\_\_\_\_ death, he was still regarded with \_\_\_\_\_. Natural religion was considered to be \_\_\_\_\_ same as atheism. This was hardly \_\_\_\_\_, since the clearly stated purpose of \_\_\_\_\_ book was to fight against atheism, \_\_\_\_\_ was growing in France. He wanted \_\_\_\_\_ show that there was a real \_\_\_\_\_ valuable natural religion available to replace \_\_\_\_\_ older religion of the Bible, which \_\_\_\_\_ Frenchmen found they could no longer \_\_\_\_\_ in. He repeatedly stated his belief \_\_\_\_\_ "one god" the creator. When Paine expressed his belief in independence, going with the trend of time, he was praised by the public; when he expressed his belief in natural religion, going against common beliefs, he was damned.

