Study of the Cloze procedure, its history and use as a measure of comprehension and readability

Eugene Janssen

Follow this and additional works at: https://digitalcommons.stritch.edu/etd

Part of the Education Commons

Recommended Citation

https://digitalcommons.stritch.edu/etd/825
A STUDY OF THE CLOZE PROCEDURE, ITS HISTORY
AND USE AS A MEASURE OF COMPREHENSION AND READABILITY

by

Eugene Janssen

A RESEARCH PAPER
SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS IN EDUCATION
(READING SPECIALIST)
AT CARDINAL STRITCH COLLEGE
Milwaukee, Wisconsin
1976
This research paper has been
approved for the Graduate Committee
of Cardinal Stritch College by

Sister Marie Collette
Advisor

Date April 9, 1975
ACKNOWLEDGMENTS

The writer of this paper would like to express appreciation to the faculty of the Reading Department of Cardinal Stritch College for their assistance in the preparation of this paper.

Sincere appreciation and gratitude is extended to Sister Marie Colette, O.S.F., who with patience, perception, and understanding, assisted in the preparation of this paper.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>INTRODUCTION.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Statement of the Problem.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Definition of Cloze</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Objectives.</td>
<td>4</td>
</tr>
<tr>
<td>II.</td>
<td>HISTORY OF THE CLOZE.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Cloze as a Measure of Comprehension</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Cloze as a Measure of Readability</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>32</td>
</tr>
<tr>
<td>III.</td>
<td>CLOZE AS AN INSTRUCTIONAL INSTRUMENT.</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Construction of the Cloze Test.</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Use of the Cloze Test</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Scoring the Cloze Test.</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Cloze in Published Textbooks.</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Summary and Conclusion.</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>SELECTED BIBLIOGRAPHY.</td>
<td>46</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Summary of the Results of Fourteen Studies Using Cloze as a Measure of Comprehension and Readability ................. 8
CHAPTER I

INTRODUCTION

This study was concerned with the cloze procedure and its use as a testing instrument in the reading program. The writer reviewed the literature concerning the cloze as a measure of comprehension and also as a measure of readability.

Reading is a skill which has been defined in a number of ways, however, this author considers reading as a thinking process involving the perception of graphic symbols and relating these symbols to past experiences to derive meaning from the printed page. Reading involves the physical act of perceiving symbols, and the mental act of deriving meaning from these symbols.

Comprehension is difficult to define. It is possible to state what comprehension is not, but some experts believe there is no adequate definition.

Yoakam describes comprehension as:

Comprehending reading matter involves the correct association of meanings with the word symbols, the evaluation of meanings which are suggested in context, the selection of the correct meaning, the organization
of ideas as they are read, the retention of these ideas, and their use in some present or future activity.¹

Edwards explains that

... continuous development toward greater reading proficiency is a process with many phases, the goal of which is comprehension of ideas. Success in the process depends on adequate motivation, a substantial background of concepts, word-perception skills, and the ability to reason one's way through smaller idea elements and to grasp, as a whole, the meaning of a larger unitary idea.²

Comprehension involves a number of abilities. Too often comprehension is thought of as the recognition of the printed word, but this word recognition is only one of the abilities necessary for comprehension. It involves definition, retention of ideas, and the use of these ideas either in the present time or at some future date.

Reading is more than the physical act of word recognition, more than verbalism. It is comprehension that makes reading more than this physical act of seeing and saying the printed word. Most authors will agree that if students are expected to understand and comprehend material used in the classroom that this material should be on an instructional level commensurate with the students' ability.

The student must be able to read the material; it must be at the grade level the student can handle. This

²Ibid., p. 213.
readability depends upon the nature of the material to be
read and the ability of the reader. To say that a particu-
lar book is readable for a student at his grade level, some
decisions must be made concerning the depth of understanding
required. What is readable depends upon the background,
abilities, and interests of the reader. There are many
ways to determine readability. Teachers are always
searching for a method that is both time-saving and accurate,
and one that the teachers can administer and score.

Statement of the Problem

The intent of this paper was to review the literature
concerning the history of the cloze procedure and its use
and development as a testing instrument in the reading pro-
gram. Much has been written, and much research has been
done concerning the cloze procedure and its use in the
elementary school. The writer of this paper will consider
the application of the cloze procedure and its adaptability
to the secondary schools.

Definition of Cloze

The cloze procedure has been defined by a number of
authors:

Cloze--a method of intercepting a message from a
transmitter [writer or speaker], mutilating its language
patterns by deleting parts, and so administering it to
receivers [readers and listeners] that their attempts to
make the pattern whole again potentially yield a considerable number of cloze units.3

The term cloze procedure refers to a reading passage in which words are systematically deleted, words that the student is then instructed to replace.4

A simplified description of the cloze procedure includes these steps:

a. passages are selected from the material whose difficulty is being evaluated.
b. every fifth word is deleted and replaced by underlined blanks of a standard length.
c. the tests are duplicated and given, and within time limits, to students who have not read the passage from which the tests were made.
d. the students are instructed to write in each blank the word they think was deleted.
e. responses are scored correct when they exactly match the words deleted.5

These definitions all seem to agree on the basic principles that the cloze procedure involves the deletion of a determined number of words which must be replaced by the student. Although the viewpoints may vary as to the strictness of the substitution of words, and as to the deletion of every fifth, tenth, or whatever word, the basic definitions appear to be similar.

Objectives

The purpose of this paper was to review the literature of the cloze procedure within the following areas:


1. the history of the cloze procedure.
2. the development of the cloze procedure as a measure of reading comprehension.
3. the development of the cloze procedure as a measure of readability.
CHAPTER II

HISTORY OF THE CLOZE

In 1953, Wilson Taylor published the first comprehensive statement of the cloze procedure. The central core of this procedure is a unit of measurement called a cloze, a term derived from closure. Gestalt psychologists applied the term closure to the human tendency to complete a similar but not quite finished pattern. This incomplete pattern is completed mentally. For example, an incomplete circle is mentally completed because its shape is familiar to the viewer.

This same idea concerning completeness can also be applied to language. Passages are mutiliated and the person tested by this procedure is asked to complete the sentence meaning as a whole. The pattern must be completed; words must be supplied that will be appropriate to the meaning of the pattern as an entire unit. As stated earlier in the definition, in the cloze test the reader decides from the context that remains what the missing words should be.

Cloze is a process in which words are deleted from the printed passage. The student is then asked to predict the exact word that was deleted and to write it in the
space provided. The ability to replace the missing words will depend upon the students' prior knowledge, general understanding of the material, context clues, knowledge of word usage. 1

In constructing a cloze test, a passage is selected from reading material that is being used in the classroom. The teacher decides on the count to be used in deleting words. For example, every fifth word may be deleted for simple narrative material, or every tenth word deleted for textbook materials. Normally, the first and last sentences are printed free from any deletions. The exercise is then typed out with the proper deletions, and with a blank space of the same length for each deletion. Clues may be given, the first letter of each word to be supplied may be given, or the length of line may vary depending upon the length of the word that is to be supplied. The student is rated on his ability to predict the word that was deleted and to supply that word in the incomplete sentence. This word must fit the context of the sentence as well as the context of the passage.

Table 1 summarizes the results of the fourteen studies that were reviewed by this author. These studies concerned the use of the cloze procedure as a measure of reading comprehension and of readability.

---

<table>
<thead>
<tr>
<th>Author</th>
<th>Population tested</th>
<th>Type of deletion</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heitzman &amp; Bloomer 1962</td>
<td>ninth grade</td>
<td>random number</td>
<td>No significant difference in reading comprehension</td>
</tr>
<tr>
<td>Bloomer 1962</td>
<td>college freshmen</td>
<td>every tenth word</td>
<td>Significant differences in reading comprehension between experimental group and the traditional remedial reading class and the non-treatment group</td>
</tr>
<tr>
<td>Martin 1968</td>
<td>college freshmen</td>
<td>lexical elements</td>
<td>No significant difference between students using cloze and those studying transformational grammar</td>
</tr>
<tr>
<td>Guice 1969</td>
<td>college freshmen</td>
<td>every fifth word</td>
<td>Significant differences for experimental group over the group receiving regular instruction in reading</td>
</tr>
<tr>
<td>Louthan 1965</td>
<td>junior high</td>
<td>nouns, verbs, adverbs, adjectives</td>
<td>Varying results</td>
</tr>
<tr>
<td>Ellington 1972</td>
<td>eleventh grade</td>
<td>tenth word</td>
<td>No significant difference between the groups in ability to comprehend material</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Grade</td>
<td>Interval</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Geyer</td>
<td>1970</td>
<td>junior high</td>
<td>not specified</td>
</tr>
<tr>
<td>Beard</td>
<td>1967</td>
<td>tenth grade</td>
<td>every fifth word</td>
</tr>
<tr>
<td>Jenkinson</td>
<td>1957</td>
<td>high school</td>
<td>every fifth word</td>
</tr>
<tr>
<td>Taylor</td>
<td>1954</td>
<td>four groups of 18 subjects each for a total of 72</td>
<td>seventh word</td>
</tr>
<tr>
<td>Bormuth</td>
<td>1962</td>
<td>fourth and fifth grade</td>
<td>every fifth word</td>
</tr>
<tr>
<td>Bormuth</td>
<td>1962</td>
<td>not specified</td>
<td>every fifth word</td>
</tr>
<tr>
<td>Alexander</td>
<td>1969</td>
<td>intermediate grades</td>
<td>every fifth word</td>
</tr>
<tr>
<td>Jefferson</td>
<td>1970</td>
<td>eighth to twelfth grade</td>
<td>nth word, lexical unit, structural unit</td>
</tr>
</tbody>
</table>
Cloze as a Measure of Comprehension

In 1966, Heitman and Bloomer conducted the Cooperative Research Project in the study of the effectiveness of the cloze procedure. Ninth graders who were involved in this study worked two cloze exercises a week for a twelve-week period. The following conditions were developed: fifth and sixth grade reading level passages were used containing a complete episode of 600-700 words. The treatment consisted of several items:

- no deletions
- random deletions
- deletions of nouns, verbs, modifiers, function words, and noun determiners.

Ten percent of the passages were deleted, each was followed by twelve comprehension questions and tests were administered in blocks of six.

Pre-tests given were the Differential Aptitude Test and the California Test of Mental Maturity. The language and reading parts of the Iowa Test of Basic Skills was given as a post-test along with a 1000 word cloze test and a comprehension test of multiple-choice items. The basic finding was that:

The use of non-overt reinforced cloze procedure does not increase reading ability during the process or as a function of post-treatment testing.

---


3 Ibid., p. 12.
No significant improvement in comprehension test scores was evident in the students, but the number of correct closures was related to the Iowa Test of Basic Skills reading scores, slightly related to reading comprehension and inversely related to recall comprehension questions on the post-test scores.

The authors felt that the value of the cloze is directly related to the method by which it is presented. They suggest that additional reinforcement of correct responses by the teacher and a plan whereby the movement through the tests depends on the quality of the responses would enhance the effectiveness of the cloze procedure in teaching.4

In 1962, Bloomer used the cloze procedure with a group of college students who were in a remedial reading program. Three groups were formed. The first received the cloze exercise, a second followed the traditional remedial reading exercise, and the third received no treatment at all. Every tenth word was deleted on material that was graduated in difficulty. Ninety-six percent correct was set as a criterion for advancement from one level to the next, and synonyms were also counted as correct by the instructor.

On the results of the Diagnostic Reading Test given as a pre-test and post-test, the group using cloze made

increases in comprehension and total reading ability over the traditional remedial class and the group receiving no special treatment.

However, there were points that can be questioned about the results of this study. No deliberate teaching was done; the instructor was merely a record-keeper. A question remains as to the lasting effects since some completed the program in only twelve sessions. 5

Martin, in 1968, investigated the effect of instruction in transformational grammar and the completion of the cloze. One hundred college freshmen were involved in this study with one half studying transformational grammar by Jacobs and Rosenbaum, and the other half receiving eighteen cloze exercises, and the control group of forty-two students enrolled in a freshman English composition course.

Lexical elements were deleted. Multiple-choice elements based on the grammatical and semantic criteria were given for each cloze blank. These multiple-choice alternatives were not used after the first week, but the cloze training consisted of teacher-led discussions upon completion of the cloze tests.

After nine weeks the tests indicated that both of the experimental groups made significant gains on the Iowa Silent Reading Test. There were no significant differences between the experimental groups. The conclusion reached by Martin was that the experimental treatments were about equal, and verbalizing the reasons for closures did seem appropriate for instruction in this method. 6

Guice, using college freshmen for his study in 1969, attempted to show that a student receiving regular instruction in reading comprehension plus use of the cloze procedure would make greater gains in comprehension when compared with students receiving only the regular comprehension instruction. Sixteen cloze exercises were used with deletion of concept words. Credit was given for synonyms as well as exact replacement. After the battery of pre-tests and the period of instruction, the post-test showed a significant gain for the experimental group over the control group.

However, as impressive as this study might seem, there are a few questions left unanswered. The number of deletions, the length of passages, and the rationale for accepting synonyms were not given consideration. Moreover, no mention was made as to whether any real instruction took place, and above all, the course content and passages used

were based on a content analysis of the comprehension test used. The author was teaching for the test.\(^7\)

Louthan, in 1965, conducted a study using the cloze test to determine the contributions to the meaning of the passage of certain kinds of words and the relative difficulty of the reading matter in a prose selection. To measure the effect of deleting certain grammatical classes of words twenty-four passages of 500-600 words were selected and prepared in seven cloze tests. All forms of the test were followed by twelve comprehension questions. Passages were both fictional and non-fiction and unlikely to appear in the school textbooks.

Cloze tests were made by dividing the passages into ten-word segments. The word closest to the end of each segment which fitted the desired grammatical class was deleted. A ten percent deletion was achieved and evenly dispersed throughout the passage.

There were seven different types of deletions used: deletion of the last word of each ten-word segment; deletion of proper and common nouns; deletion of specific verbs except function verbs and auxiliary uses; deletion of specific modifiers, adjectives, and adverbs; deletion of prepositions and conjunctions; deletion of function words; deletion of the substantive use of pronouns.

\(^1\)Billy M. Guice, "The Use of Cloze Procedure for Improving Reading Comprehension of College Students," *Journal of Reading Behavior* 1 (Summer 1969):81-93.
Pupils in a junior high school in New York were selected at random and assigned cloze test forms. Each of the 236 students was given only one form and would receive all six experimental exercises in one particular type of deletion test. Students were asked to complete the cloze test and to answer the comprehension questions without referring to the reading passage. The experiment lasted for three school weeks. Data provided were the number of correct responses to the comprehension questions following each of the cloze tests.

The results of this study would seem to indicate deficiencies in the ability of students to answer comprehension questions subsequent to the presentation of reading material by the cloze procedure. The control group, those students whose passage had no deletions, seem to be superior to some groups using cloze procedure material with a particular kind of deletion. The comprehension exercises of the control group were superior to those of the random deletion group, the noun deletion group, the verb deletion group, and the modifier deletion group. The last three classes carry the meaning of the written material and when one word out of ten is deleted there is a loss in comprehension. If prepositions and conjunctions were deleted there was no difference between the two groups and the control group. The deletion of noun determiners produced a gain in comprehension and the scores were higher than those following any other form.
Several conclusions can be drawn on the basis of the results of the study. Removal of function words probably tends to draw attention to the larger units of meaning of the passage. While the deletion of words like nouns, verbs, and specific modifiers draws the attention of the reader to units of less meaning.

The exact replacement of a deletion may cause the comprehension to suffer because the reader is not sure of the response and becomes frustrated.

Repeated reading of a passage for the purpose of supplying the missing words would appear to be a form of reinforcement. Repeated reading of material that is incomprehensible might reinforce the frustration factor which indicates that the materials are beyond the readers' grasp.

The time required to supply the deleted words will also increase exposure of the reader to the material. This exposure will either enable the reader to crystallize information by replacing words, or might allow the reader time to realize that there are gaps in his understanding of the material.

Systematic deletion of all of the words called structural function words will probably produce a marked loss in reading comprehension. This study demonstrates that an efficient reading process may be induced in students.
Also, the understanding of what kinds of words bear the burden of communication will draw the reader's attention to words of those classes. 8

In 1972, Ellington conducted a study that compared the scores of three groups of eleventh grade students on the Cooperative English Test: Reading Comprehension, with a group doing cloze reading comprehension exercises, with a group receiving conventional reading exercises, and with a group receiving no reading comprehension exercises.

The eleventh grade students were randomly selected for one of the three groups. The cloze tests consisted of 75 cloze comprehension exercises which deleted every tenth word. The instructional material used was the Power Builders for grades eight to twelve in the Science Research Associates Laboratory, IIIa. The conventional reading class used the exercises from the same Power Builders, and the third group received regular English instruction. The investigator presented all of the material to the cloze and conventional group whereas the regular English class was conducted by the English teacher. The cloze and conventional groups were expected to work for fifty-five minutes each day for a period of six weeks.

Analysis of the data revealed no significant differences between the three groups.  

Geyer investigated the effectiveness of the cloze procedure as a predictor of students' abilities to comprehend social studies material. Three hypotheses were considered:

1. The cloze score can predict the ability to comprehend social studies material better than standardized reading test scores.
2. Cloze test scores are more reliable in predicting comprehension of social studies than I.Q. scores.
3. Cloze tests are a better predictor of comprehension of social studies material than the previous social studies grades.

Two junior high school history texts were used with 201 junior high students. The data used includes the pre-reading cloze, an I.Q. test, previous social studies grades, a standardized reading test, and the criterion measure, a multiple-choice test. All participants completed one of two cloze test forms and then read the chapter in the textbook. After taking the cloze test the students read the chapter and completed the multiple-choice test. The following conclusions were drawn:

---

1. The cloze test cannot predict students' abilities to comprehend social studies material better than the standardized reading test.

2. Cloze test scores are not more reliable in predicting comprehension of social studies than I.Q. scores.

3. Cloze tests are not a better predictor of comprehension of social studies material than previous social studies grades.

It was further concluded that the cloze test was not the best predictor of ability to comprehend specific instructional material but it was correlated highly enough to warrant further investigation.10

In 1967, Beard conducted a study using 250 tenth grade students to investigate differences in the comprehensibility of textbooks used in content area studies. This grade level was chosen because these students would have had similar instruction in junior high school. Students were divided according to ability. Fifty students were selected for each percentile range 1-19, 20-39, 40-59, 60-79, and 80-99 on the verbal section of the School and College

Ability Test. Cloze forms were assigned to each ability level so that each cloze testing procedure was replicated for each of the five ability groups.

Passages were selected from the main body of each of ten textbooks chosen. Pages were selected using a table of random numbers. Passages following subtitles were used if they did not duplicate a previously chosen topic. Each passage contained 250 words of continuous prose, and these 250 words did not contain more than ten quoted words.

Using a systematic word deletion process, every fifth word was replaced by a blank of uniform length. Deletion began with the first word in the first sentence and other forms were prepared by deleting the second, third, fourth, and the fifth word. This process confirmed the fact that every word was included as an item to be replaced. Only exact replacements were counted as correct, and a score for each passage was obtained by averaging the percentage of words correctly replaced to its cloze test forms.

Results of the cloze testing were summarized for each type of content by computing means and standard deviations for the cloze scores and for the expressional elements of each content. The results show that the average level of difficulty of textbooks does not differ significantly
among the subject matter content areas. The mean cloze scores for world history, American government, and biology were similar. The mean score for chemistry was higher and these scores were less variable. One reason for this fact might be that it was difficult to find a continuous passage of prose in the chemistry books since the passages were interspersed with formulas in the text.

Small and insignificant differences among content were found for the means of the comprehensibility measures. The conclusion reached was that the difficulty of prose used in current high school textbooks is about the same for American government, world history, biology, and chemistry. 11

From the rather sketchy details that this writer was able to find, Jenkinson conducted a study using the cloze test as a diagnostic instrument. After giving the cloze test to high school students, the students with high and low scores were selected for interviews. During this interview another cloze test was given but this time each student was asked to verbalize his reason for his choice of words. This procedure was found to be useful as a diagnostic instrument in the reading clinic. This method

does provide reinforcement which is necessary for good learning plus it provides for interest and growth in vocabulary improvement. 12

Cloze as a Measure of Readability

Teachers are constantly questioning the readability of many of the textbooks that students encounter in daily work. Material that is too difficult can harm reading progress if work in this material is maintained over a period of time, whereas easy material can have similar effects if taken as a steady diet of reading. Since the difficulty levels of some materials are not often known, the teacher is placed in a difficult situation.

Several formulas have been devised to serve as a measure of readability. The Flesch and Dale-Chall are perhaps the two most widely-known and commonly used. Both formulas require the counting of words in the passage. The Flesch method counts the number of syllables per word and the number of words in a sentence. Dale-Chall counts words in a sentence and the number of words not found on a list of commonly used words. Both of these techniques are based upon the assumption that passages with poly-syllabic or unusual words and long sentences are more difficult to read. Both of these formulas can be misleading if the passages

contain short words of an abstract nature. Neither of these two methods take into consideration the particular background of experience or interest of the reader.

In 1953, Taylor conducted a study to investigate the possibilities of using the cloze procedure as a measure of readability. Two considerations governed this investigation. Passages were selected that seemed to be distributed evenly over a long hard-to-easy range, and materials which would show that cloze procedure could deal with passages which the Flesch and Dale-Chall formulas could not handle.

Both considerations were served by passages from Caldwell, Stein, and Joyce, as well as passages from Boswell, Huxley, and James. The passage from Caldwell is by far the easiest to read of all the passages, and the passages from Stein and James are the next-hardest and hardest respectively. It was expected that the readabilities of the three selections would be under-rated or over-rated by either or both formulas. Two other passages from Swift and Dickens were added bringing the total to eight for this part of the study.

These eight passages were ranked by their median cloze scores and their order was considered a predictor of how large groups would rank them.

The chief hypotheses were:

1. The readability order of eight passages would dependably predict the rank order of cloze
scores given by other larger samples and subjects.

2. Cloze procedure would handle passages which the formulas could not.

3. Cloze prediction would be more successful than formulas in guessing the rank of final cloze scores.

4. The overall differences among the cloze scores of the passages would be found to be significant when subject to analysis of variance.

Each subject received two of the eight passages. Each passage was administered to eighteen subjects. The sixteen kinds of pairs were represented four times each in the final data to which a total of seventy-two subjects contributed.

The results would seem to indicate that the cloze procedure came closer than either formula in ranking properly the relative reading level of certain passages. Flesch ranked the Stein passage as easy reading while Dale-Chall ranked it within the comprehension of fourth or fifth grade school children. The formulas ranked the Stein passage as most readable because this passage is composed of short words and short sentences and these formulas are concerned with words and syllables. The cloze procedure more adequately ranked the readability of these three passages. It ranked the Stein passage as most difficult to read. The
passage from Joyce contained many words not commonly found in the dictionary. The formulas are only concerned with syllables and rated this passage as easy reading whereas the cloze procedure which considers context rated this passage as most difficult reading.

All the hypotheses developed in this study were accepted. The conclusions do not bear much generalization. They apply only to the specific passages employed, but they were consistent.13

Bormuth in a study conducted in 1967 provided a frame of reference that pertains to students at all levels. Even though the subjects used were fourth and fifth grade students, the intention of this study was to generalize the results to students at every grade level with the possible exception of grade one and two.

The rationale behind this study was that multiple-choice tests have been in use for a long time and have developed a frame of reference in readability and classroom practice. A student must answer 75 percent of the items over a passage before the passage is suitable for the student's use. A score between 75 and 90 percent indicates that the material is suitable for supervised study. A score above the indicated range indicates that the material is suitable for independent study, while a score below this

range indicates the material is too difficult for instruction. The ratio of items answered correctly on the test over a passage may be a measure of the readability of the passage.

A fifty item cloze test and a thirty-one item multiple-choice test were made over each of nine passages. Tests were validated by experts and given to 73 pupils in grades four and five. The applicability of this test to high school is explained by Bormuth when he contends that the variables which influence the comprehension of readers at one level of reading ability have a similar influence on readers on other levels.

The cloze tests contained about 275 words and every fifth word was deleted. Dale-Chall formulas rated the material from 4.5 to 6.5 grade level. First, the cloze tests were given. Only exact answers were counted as correct. Three days later, students were given the multiple-choice tests after having read the passages.

The major conclusions of this study indicate that if the conventional readability standards are accepted, a cloze score of 38 percent indicates the material is sufficiently understandable. The student will be able to answer 75 percent of the multiple-choice questions. Taking into consideration the elements of guessing on the multiple-choice test a cloze test score of 43 percent is necessary to reach the same criterion level.
Hence, Bormuth's contention is substantiated that the cloze tests are valuable as a method of determining how well instructional materials can be understood by the students. The formulas express the difficulty levels in terms of average grade placement, but the teacher must judge whether a student's score shows that he can understand materials at an acceptable level.14

Another study by Bormuth in 1962 investigated the validity of the cloze test as a measure of readability. The readability of a passage is defined by the amount of comprehension with which it is read. Three assumptions were involved in this study. These related to the cloze tests as a measure of comprehension, the cloze test efficiency when used to discriminate among the difficulties of tests and individual ability, and the validity of the cloze as a measure of the amount of comprehension with which the passages are read.

Nine passages were used which were divided among three classes of subject matter and three reading levels of reading difficulty. By counting a subject's correct response to the items measuring one skill, a measure of ability in that skill was obtained. Cloze tests were made over each passage by replacing every fifth word with a blank. Fifty items were obtained. Exact word replacement was counted as

correct and all subjects, three groups of fifty each, took the tests. Students were from the upper grades.

The findings support the hypothesis that the cloze tests are valid measures of comprehension ability. Scores on the cloze tests correlated with the scores on the comprehension tests which were written over each passage and measured seven comprehension skills. All correlations between cloze and the comprehension tests over the same passages were significant.

The findings supported the assumptions that cloze tests are efficient when used for the purpose of measuring readability and comprehension ability. Results also supported the hypothesis that the cloze tests are valid measures of readability.\footnote{John R. Bormuth, "Cloze Tests as a Measure of Readability and Comprehension Ability," Dissertation Abstracts 23 (May 1963): 4218-4219A.}

Alexander, in 1969, investigated the cloze procedure as a measuring device to identify the reading levels of pupils in the intermediate grades. Since Bormuth contends that the studies regarding readability are applicable to any grade this study does generalize the results to students at other grade levels. It is important for success of the students to identify the independent, instructional, and frustration reading levels. An \textit{Informal Reading Inventory}
seems to be the most efficient test of readability but because it must be administered individually the usage in the classroom is limited.

The purpose of this study was to determine that the cloze test could furnish information similar to that of the formal reading inventory, and to determine the relationship of scores from the two informal reading tests and the results from the standardized reading tests and the intelligence measure.

Three hundred sixty-five students from the intermediate grades of a suburban Chicago school were used in this study. The Lorge Thorndike Intelligence Test M Level Edition, the Gates-MacGintie Reading Tests Survey, an informal reading inventory, and a cloze test were administered to each group. The same reading passages were used in the informal reading inventory and the cloze test. Readability levels of the material was subjected to the Spache Readability Formula and the Dale-Chall Readability Formula.

From the statistical analysis it could be concluded that a measuring instrument based upon the cloze procedure can identify the reading levels with a degree of accuracy that compares to the informal reading inventory. Those cloze scores below 47 percent identified the frustration level, scores from 47 to 61 percent identified the instructional level, and over 61 percent identified the independent reading level. The cloze test scores correlated with the
scores of the standardized tests and an intelligence test as well as the correlations of the informal reading tests with the same testing instrument.

The conclusions reached verify the fact that the cloze test can be used to identify the independent, instructional, and frustration reading levels of students. The cloze percentage criteria for the three reading levels are only one set of reference points. On the basis of classroom examinations other criteria might be suggested.  

Jefferson studied the question of lexical and structural items as predictor of readability for high and low ability readers. He set down six hypotheses:

a. there are significant differences among certain linguistic variables as measured by the cloze, specifically between lexical, structural, and nth deletion cloze responses;
b. there are significant differences among subjects at different levels of reading ability as measured by the cloze;
c. there are significant differences among levels of readability as measured by the cloze;
d. there are significant interactions among certain linguistic variables as measured by cloze, and different levels of reading ability;
e. structural words or units are the best predictors of reading difficulty among the linguistic variables tested by cloze measures;

f. the structural deletion scores of poor readers are the more consistent predictors of reading difficulty than are the structural word scores of better readers. 17

Jefferson's subjects were the students in grades eight to twelve of a high school in northeast Georgia. All students took STEP 3A and 2A tests, and two weeks later were given the cloze instrument. Subjects were divided into three equal groups. The subjects at each reading level, high, middle, and low, were randomly assigned to one of nine treatment conditions within each level. These nine treatment conditions were not specified by Jefferson but his description of the procedure suggested that six passages were used and deletions were made in three separate ways: nth unit, lexical unit, and structural unit.

Results showed that there were significant differences among the nth, lexical, and structural deletions. Lexical cloze was most difficult, nth deletion cloze units were intermediate in difficulty, and structural cloze units were the easiest to replace. There were no statistically significant differences among levels of reading ability as measured by the cloze test. A significant interaction was found between deletion type and readability. The hypothesis that there are significant interactions among one or more of the linguistic variables and reading ability was not confirmed.

It is difficult to evaluate Jefferson's findings without some specific information which he does not provide. For example, the total number of subjects involved is not given, nor the criteria for "high, middle and low" groups. The passages are described as "a reasonably representative sample of high, middle and low readability" but the range of readability was not reported. Was it grades eight to twelve, or was it a wider range? Were all the passages within the capability of the "high group" or were some of them beyond reading achievement level?  

Summary

A total of fourteen studies were reviewed by the writer of this paper. Nine concerned cloze as a measure of comprehension and five studies considered cloze as a measure of readability. While many of the studies considered the cloze procedure to be an effective and accurate measure of comprehension and readability, some of the conclusions reached were not as definitive.

Heitzman and Bloomer concluded that the cloze tests do have considerable value, but this value is related to the method by which it is used. Additional reinforcement would make the test more effective.

Jenkinson conducted a study in which students were asked to make a response to the cloze test and then were

\[18\] Ibid., pp. 172-178.
given the opportunity to justify their answers. This method would provide for the reinforcement that is necessary to good learning, and would serve as a diagnostic instrument.

Bloomer's independent study again reached the conclusion that the cloze procedure could produce good results in a remedial reading program. However, there are a couple of questions in regard to the study. There is no mention concerning any actual instruction that took place. Some students completed the work in twelve sessions which does not appear to be long enough to have any lasting effects.

Martin's investigation concerned the deletion of lexical elements. He concluded that there were significant differences between the experimental groups, but there was no significant differences between the group that used the cloze procedure and the group that studied transformational grammar. Martin favors the verbalizing of reasons for answers to the cloze test as one method of instruction.

The study conducted by Guice showed significant gains for the experimental group which used the cloze procedure. But as indicated there are a number of questions about this study. No mention was made as to the number of deletions, the length of passages used in the study, and no reason was given for accepting synonyms. It would seem that the author was teaching for the test since passages were based on the comprehension test used.
Louthan conducted a study attempting to determine the contributions of certain kinds of words to the meaning of the passage. This study deleted grammatical classes of words. Deletion of grammatical elements does draw the attention of the reader to the words that bear the burden of communication.

Ellington's study showed results contrary to the other results presented. The conclusion reached was that there was no significant difference in reading comprehension between the groups used for this study. It does appear rather difficult to accept the fact that the two groups that were concerned with reading comprehension did not make any significant improvement over the regular English class.

Geyer's study concluded that the cloze test cannot predict the students' abilities to comprehend social studies any better than standardized reading tests. The cloze is not more reliable than the I.Q. in predicting comprehension of social studies material, and the cloze is not a better predictor of comprehension than previous social studies grades. However, Geyer concluded that the correlation between cloze scores and comprehension scores was high enough to warrant further study.

Beard was concerned with the comprehensibility of the textbooks used in some of the classes in the content areas. No mention was made as to the series of textbooks considered but the results indicate that there is no significant differences in the difficulty of the prose used in the current
high school textbooks in American government, world history, biology, and chemistry.

Taylor's study regarding readability was detailed. This study used actual passages from books that are considered excellent reading by many people. This study pointed out the differences between the readability scores obtained by the formulas as opposed to readability scores obtained by the cloze procedure. The conclusion reached was that the cloze procedure more adequately ranked the readability of the passages considered than the formulas.

Bormuth used the multiple-choice test and the cloze test as a measure of readability. Both types of tests were given to students and the results of these tests were compared to determine what percentage of cloze elements must be correctly scored to compare to a similar score on the multiple-choice test.

The second study by Bormuth used the cloze as a measure of readability and comprehension. Results of this study indicate that the cloze tests are valid measures of comprehension. Correlations between cloze and the results obtained on teacher-made comprehension tests were significant.

Alexander used the cloze procedure to measure the reading levels of pupils in intermediate grades. Scores from the cloze tests were compared with the scores achieved on other reading tests and the results indicate that the cloze can be used to identify reading levels.
The writer suspends judgment concerning Jefferson's conclusions because of unreported information about his procedure.
CHAPTER III

CLOZE AS AN INSTRUCTIONAL INSTRUMENT

Since the cloze procedure was first introduced by Taylor, in 1953, there have been numerous studies made of its usefulness as a teaching instrument. Many of these studies were not reviewed by the present paper because they were concerned with elementary students. The writer reviewed studies that involved high and college students in the previous chapter.

The cloze procedure is one way to take a measure of language between the reader and the writer. The reader is asked to replace the exact words that have been deleted from the text under study. If the reader and writer have similar interests and experience the reader should be able to accurately anticipate the exact words that have been deleted. Exact word replacements are indicative of the reader's ability to understand the meaning of the passage, context clues, and word usage.

Construction of the Cloze Test

The method of construction of a cloze test is not very difficult. The teacher selects a passage from the material that is being used in the classroom. After deciding
upon the number of deletions, every fifth or tenth word, the exercise is typed for distribution to the students. The first sentence is typed without any deletions. Each deletion is represented by a blank of uniform length, usually fifteen spaces, and the last sentence is also typed without any deletions. Students taking the test are asked to fill in the blanks with the exact word. The words must be appropriate to the context of the passage.

Cloze tests can be prepared in a variety of ways:

--lexical items, nouns, verbs, and adjectives can be deleted to emphasize the elements of language.
--parts of words can be deleted, all but the suffix, the prefix, or initial sounds can be deleted to emphasize word recognition.
--prepare a cloze test with the first or last word of the sentence deleted to focus instruction on syntax of the language.
--delete items and ask for synonyms for substitutions in order to work on vocabulary improvement.
--use different passages from various authors to show variation in language or style.
--preparation of tests with the deletion of certain sounds, for instance, the long a sound, or the short e sound, or even the blends bl, gn, or whatever. This will encourage word analysis. ¹

Use of the Cloze Test

The teacher is often faced with the task of placing students in appropriate reading materials. Materials should be on a suitable reading level of difficulty. These materials should be difficult enough to permit practice in the reading

skills but not too difficult as to develop faulty reading habits, or a sense of frustration. One of the standard procedures used in testing readability of material is to have the student read orally a sample passage from the material being considered, and then the student is expected to answer questions about what was read. This method requires a great deal of time to administer, and careful and special training for the examiner. Because of these two restrictions this type of test is not used very often.

The cloze procedure can determine readability of the materials quickly and reliably. Several sets of tests can be compiled from the material under consideration with several passages from different sections of the book being used as a test of readability. This random selection of passages from the textbook should be representative of the reading difficulty of the material. Each passage used should be of continuous text of approximately 250 to 300 words in length.

Bormuth gives a step-by-step description of the process:

a. Select several passages from a book. Each passage should begin at the normal beginning of a paragraph and should be at least 250 words in length. The test should contain exactly fifty items.

b. The tests are then given to students at the grade level at which the text is most commonly used. Twenty-five to thirty students are usually enough to obtain reliable results.

c. The mean score of each test is calculated and then the mean of each test is calculated. Following this the mean of the mean scores is calculated.

d. Finally, the test whose mean score is closest to the mean of the entire set of tests is selected and the rest of the tests are discarded.
e. When a test has been selected for each of the texts a teacher is likely to use, the tests can be mimeographed and compiled into booklets which can be administered as group tests. When a student's score falls between 44 and 57 percent on one of these tests, the materials are at the level of difficulty thought to be suitable for use in supervised instruction. Above 57 percent indicates material suitable for independent study.²

If, after taking the test, the correct answers are given to the students they should have immediate reinforcement of their responses. This will provide a teaching situation somewhat like programmed instruction.

Cloze tests can be used to increase vocabulary. Synonyms can be accepted in tests given as exercises for teaching. However, the teacher should insist that students justify their response which will aid the teacher in increasing the students' familiarity with words in different situations.

Culhane states that by using the cloze procedure which deletes nouns and verbs a greater understanding of factual material can accrue. Deletion of pronouns could also be used as a method of teaching correct use of the pronoun.³

In administering a cloze test the teacher first decides upon the rate of deletion of words in the passage


involved. Students are instructed to read the entire passage before they begin the cloze test. This first reading enables the students to make the most of contextual clues and information when they try to fill in the cloze test blanks. This pre-reading will also encourage the students to develop the habit of reading the material before and after the deletion process. Students should be taught the value of context clues as a word recognition strategy.

The passages can be read aloud by the student or by the teacher, and suggestions as to appropriate words to use can be offered by the students. All acceptable responses are taken but students are expected to give reasons for their word selection. Through this procedure the teacher encourages understanding of the language and the teacher becomes aware of the particular needs of the students. For example, if the student replaces verbs instead of the required nouns the teacher is made aware of the need for additional instruction in this particular area.

After the cloze test has been completed the answers offered by the students are compared to the original test. Discussion should follow as to why some words were acceptable and others were not, and how the meaning of the entire passage is affected by certain responses. The use of "bike" for "bicycle" does not affect the meaning of the passage, but the substitution of "boat" for "book" would certainly change the meaning of the passage. The teacher should stress the use of context clues in the selection of
words. These clues can come before or after the blanks, in the beginning, in the middle, or at the end of the passage. These clues that the student is encouraged to look for and to use give the student skills that can be used in independent study.

**Scoring the Cloze Test**

Exact word replacement is scored as correct. Synonym replacement would make the scoring very difficult and necessitate many decisions as to the value of one synonym over another. Each exact replacement is counted as one point and is used to determine percentage scores.

The number of blanks completed with the exact words are counted as correct. If a student completes 45 to 57 percent of the blanks correctly the material is probably on his instructional level. Fifty-seven to 60 percent correct completion represents the independent reading level. Completion scores below 45 percent would represent the frustration level. However, formulas have been developed whereby these figures can be used to determine how the cloze procedure corresponds to the various reading levels established on an Individual Reading Inventory. The number of exact replacements is divided by the total number of blanks. This answer is then multiplied by 1.67 to determine average comprehension. For example:
These comprehension scores can then be related to the reading levels. The most widely used criteria for determining reading levels were developed by Betts. These criteria place the independent level at 95 percent or better comprehension, the instructional level at 75 percent or better comprehension, and the frustration level at less than 50 percent accurate in comprehension ability.5

**Cloze in Published Textbooks**

This author reviewed at least fifteen series of textbooks, many of which are currently in use in the schools today, and discovered only two instances where the cloze was used as an exercise in the text.

**Tactics in Reading A** used the cloze procedure but in a slightly variant form. The exercise was given to the students after the assigned reading was completed so it was not a true cloze test as considered in this paper.6

---


6**Tactics in Reading A** (Glenview, Ill.: Scott Foresman and Company, 1972), p. 196.
Sweet, author of *Artes Latinae*, uses the cloze procedure in the reference notebook that accompanies the text. Here again the exercises are to be completed after studying the reading passage. The student is asked to complete the closure in a variety of ways: beginning syllables, ending syllables, and completion of endings, all regarding the same passage. Upon completion of the notebook the student has done at least six cloze exercises on each particular passage.

**Summary and Conclusion**

Results from the fourteen studies surveyed are not in complete accord. It would appear that while the majority of studies supported the conclusions regarding significant differences in comprehension and readability produced by the cloze procedure, other studies arrived at the opposite conclusion. The enthusiasm and results produced by one favorable study was often softened by the report of opposing conclusions of another researcher. This divergent thinking should foster an attitude of inquiry and from this inquiry the reader will have to form his own opinion supported by data from studies conducted in the use of the cloze.

---

This searching and questioning of the value of the cloze should result in more studies concerning the cloze procedure as a testing instrument. Certainly the number of studies available involving elementary students is greater than those involving secondary students. This would seem to indicate that further study concerning the cloze procedure as it applies to secondary schools must be completed in order to arrive at a more definitive conclusion.
SELECTED BIBLIOGRAPHY
BIBLIOGRAPHY

Books


Tactics in Reading A. Glenview, Ill.: Scott Foresman and Company, 1972.

Articles in Journals


