

Lindenwood University

Digital Commons@Lindenwood University

---

Theses

Theses & Dissertations

---

8-1980

## An Investigation of a Matching Lexical Cloze Procedure with Eighth Grade Students

Elaine S. Bewig

Follow this and additional works at: <https://digitalcommons.lindenwood.edu/theses>



Part of the Education Commons

---

1980

AN INVESTIGATION OF A MATCHING  
LEXICAL CLOZE PROCEDURE  
WITH EIGHTH GRADE STUDENTS

*[Faint signature]*

Education Degree

BY *[Faint signature]*

ELAINE S. BEWIG

Submitted in partial fulfillment of the requirements  
for the Master of Arts in Education degree  
The Lindenwood Colleges  
August, 1980



Thesis  
B4681  
1980

Accepted by the faculty of the Department of Education, The Lindenwood Colleges, in partial fulfillment of the requirements for the Master of Arts in Education degree.

Shirley D. Rasmussen  
Faculty Tutor

Education Faculty

Jeanne Donovan

provides a good group estimate of a student's instructional  
reading level and a means to use as a diagnostic, weekly,  
series of tests for assessing student ability to read  
and instructional text ABSTRACT

In order to systematize the item writing procedure for developing reading comprehension measures for the regular classroom, the major purpose of this study was to compare the reliability, validity, and difficulty of a matching lexical cloze test, two traditional or passage-question group informal reading tests, and a standardized measure of reading comprehension, the Stanford Achievement Test. One of the traditional group informals was composed of two reading sections, each followed by 20-25 literal comprehension questions which were developed by Finn in his dissertation of 1973. The other traditional group informal was composed of two reading selections taken from a Scott-Foresman reading series followed by 20 questions based on Bormuth's Wh-rote question format developed in 1970. The matching lexical cloze tests deleted only nouns, verbs, and verbals and the correct choices were listed at the bottom of each 100-125 word selection along with six distractors for each passage. Thirty-five eighth grade students took all four types of silent reading tests. A language arts teacher also ranked the reading ability of these students and judged their ability to read the informal passages. Results indicated that the matching lexical cloze test

provides a good group estimate of a student's instructional reading level and appears to be an objective, easily developed method for assessing student ability to comprehend instructional text materials.

CHAPTER

I.	STATEMENT OF THE PROBLEM . . . . .	1
	Introduction . . . . .	1
	Justification for the Study . . . . .	2
	Definition of Terms . . . . .	3
	Research Hypothesis . . . . .	4
	Limitations . . . . .	5
II.	REVIEW OF RELATED LITERATURE . . . . .	17
	Non-Traditional Reading Methods . . . . .	17
	Suspension of Class Reading List	
	Standardized Tests . . . . .	18
	Objectivity of Class Records and	
	Individual Reading Inventories . . . . .	21
	Other Effects . . . . .	25
	Summary . . . . .	27
	Conclusions . . . . .	28
	Final Study . . . . .	29
III.	METHOD AND PROCEDURES . . . . .	35
	Sample Used . . . . .	35
	Instruments . . . . .	36
	Procedure . . . . .	37
	Data Analysis . . . . .	38
IV.	ANALYSIS OF DATA AND DISCUSSION . . . . .	44
	Expenditure in Florida Book . . . . .	44
	Fluency . . . . .	44
	Reading Inventories . . . . .	45
	State . . . . .	46
	Expenditure . . . . .	46
	Expenditure . . . . .	47
	Expenditure . . . . .	48
	Expenditure . . . . .	49
	Expenditure . . . . .	50
	Expenditure . . . . .	51
	Expenditure . . . . .	52
	Expenditure . . . . .	53
	Expenditure . . . . .	54
	Summary of Findings . . . . .	55

TABLE OF CONTENTS

Chapter

I. STATEMENT OF THE PROBLEM . . . . . 1

    Introduction . . . . . 1

    Justification for the Study . . . . . 2

    Definition of Terms . . . . . 7

    Research Hypothesis . . . . . 9

    Limitations . . . . . 9

II. REVIEW OF RELATED LITERATURE . . . . . 12

    Non-Traditional Reading Measures . . . . . 12

    Comparison of Cloze Formats and  
    Standardized Tests . . . . . 21

    Objectivity of Cloze Formats and  
    Informal Reading Inventories . . . . . 24

    Order Effects . . . . . 27

    Summary . . . . . 27

    Conclusions . . . . . 29

    Pilot Study . . . . . 29

III. METHODS AND PROCEDURES . . . . . 35

    Sample Used . . . . . 35

    Instruments and Materials Used . . . . . 36

    Procedures . . . . . 39

    Data Analysis . . . . . 41

IV. ANALYSIS OF DATA AND FINDINGS . . . . . 44

    Research Hypothesis One . . . . . 44

    Statistical Hypothesis One . . . . . 44

    Research Hypothesis Two . . . . . 45

    Statistical Hypothesis Two . . . . . 46

    Research Hypothesis Three . . . . . 46

    Statistical Hypothesis Three . . . . . 46

    Additional Purposes One . . . . . 48

    Additional Purposes Two . . . . . 52

    Additional Findings . . . . . 55

    Summary of Findings . . . . . 56

V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS . . . . .	58
Conclusions . . . . .	61
Discussion of the Findings . . . . .	62
Recommendations from the Data . . . . .	67
Recommendations for Further Research . . . . .	67
APPENDIX A . . . . .	70
APPENDIX B . . . . .	74
APPENDIX C . . . . .	82
SELECTED BIBLIOGRAPHY . . . . .	91

## LIST OF TABLES

1. Correlations of Total Scores on Non-Traditional Tests . . . . .	45
2. Correlations of Total Scores on Non-Traditional Tests and the Stanford Achievement Reading Comprehension Subtest . . . . .	47
3. Ranked Correlations on the Matching Lexical Cloze and Teacher Rankings . . . . .	48
4. Regression Ranges and Scores . . . . .	50
5. Visual Inspection (Averages) MLC and IRI Scores at the Same Levels . . . . .	51
6. MLC Accuracy of Prediction of Reading Level (MLC Scores Compared with Teacher Ratings) . . . . .	53
7. IRI Accuracy of Prediction of Reading Level (IRI Scores Compared with Teacher Ratings) . . . . .	54
8. Reliabilities (KR-20) . . . . .	55

## CHAPTER I

### STATEMENT OF THE PROBLEM

#### Introduction

In light of the "Back to Basics" movement in education today, administrators as well as teachers have become increasingly concerned with reading levels of students, particularly reading comprehension levels. For many years texts were selected and "taught" with little or no evaluation regarding an individual student's ability to comprehend the material these texts contain. Due to this growing concern involving comprehension levels of individual students and the inadequacy of current evaluative measures for particular content areas (e.g., standardized tests), it becomes necessary for individual course instructors to develop an evaluative instrument for the specific texts they use.

Such an instrument must measure both what it claims and be consistent on a well-defined scale. It must also be easily constructed by any classroom teacher as well as provide an objective measurement of student comprehension. However, many researchers have criticized the objectivity of test developers as well as the replicability of various test designs. Thus,

most individually designed tests are questionable in this area.<sup>1</sup>

It is generally agreed that conventional cloze test procedures are more objective in nature than evaluative instruments which require the writing and interpretation of questions and responses. In addition, Bormuth has been able to determine independent, instructional, and frustration levels for these conventional cloze tests.<sup>2</sup> It has been found that students feel this type of test design is anxiety producing<sup>3</sup>; nonetheless, these tests are more objective in nature than other non-traditional or conventional measures.

Therefore, it becomes necessary to develop some type of evaluative measure which individual teachers will be able to construct that will objectively measure reading comprehension, will systematize the item-writing process, and will provide a minimal amount of anxiety for students. This type of instrument is generally termed a non-traditional measure of reading comprehension.

### Justification for the Study

#### Cloze

The conventional cloze developed by Taylor and popularized by Bormuth<sup>4</sup> is the most widely used non-traditional measure of reading comprehension. The conventional cloze test deletes every fifth, or nth,

word of a passage and requires the reader to supply an exact substitution of the deleted word. This type of test instrument has undergone much research in the last twenty-five years which has produced the following conclusions:

1. The cloze technique correlates well with measures of general comprehension as indicated by standardized reading tests.
2. The cloze test correlates better with specific teacher-made or subjective multiple-choice measures of literal comprehension over the same materials than general standardized comprehension tests.
3. The cloze technique can be replicated without bias of the test maker.
4. The cloze technique ranks the readability of passages more effectively than traditional readability formulas.
5. The cloze technique discriminates among individual reading abilities at different levels of achievement.
6. The cloze test allows one to test the difficulty of various word forms, linguistic structures, and sentences within a passage.<sup>5</sup>

This test can be developed and scored quickly, is easily replicated, eliminates the subjectivity of item writers and the ability of students to answer items without reading the passage, and removes inconsistencies between the language of the passage and the language of the questions.<sup>6</sup>

The primary disadvantages of the cloze test are:

1. Student anxiety is heightened due to unfamiliarity of format.
2. Exact deleted words are the only correct responses. (i.e. Suitable synonyms are non-acceptable responses.)
3. It measures more of the syntactic than semantic component of reading comprehension.
4. It must be hand scored.
5. It may not measure information gain as a result of reading a passage.<sup>7</sup>

### Multiple-Choice Cloze Tests

The multiple-choice cloze procedure which deletes only nouns, verbs, adjectives, and adverbs was developed by several researchers. This type of instrument is less anxiety-producing to students than the conventional cloze but retains the advantage of objectivity. This measure appears to have more validity for teachers due to the fact that it is more related to the semantic aspect of comprehension and is more easily scored than the conventional cloze procedure.<sup>8</sup>

The primary disadvantage of this type of format developed by O'Reilly et al. is the necessity to consult word lists or computer banks to find suitable distractors for each item.<sup>9</sup> The time consuming nature of this operation tends to make it impractical for use by the classroom teacher.

### Matching Cloze

Several studies principally by Propst and Baldauf in 1979 and Baldauf et al. in 1980 with elementary and secondary ESL students have shown that a matching cloze technique produces reading comprehension tests that are reliable and valid and can be easily constructed by the classroom teacher. Baldauf et al. found that this type of test provided more reliable estimates of pupil performance than did standardized tests and had better validity estimates of pupil performance than did other

types of cloze tests or standardized tests when using teacher ratings as the criterion. The tests were constructed utilizing regular classroom materials and used mutually exclusive deletions with only correct responses listed in the margin for each thirty to forty word section.<sup>10</sup>

#### Matching Lexical Cloze

The matching lexical cloze format and test development used by this writer was not similar to any found in any research. The matching lexical cloze format consisted of two pages of only noun, verbals, and main verb deletions with twenty deletions per page. The correct responses along with six distractors selected from material immediately preceding the passage were listed at random at the bottom of each page. These tests contained lexical deletions, distractors, and repetitions and were administered to regular, English speaking students. Therefore, the difference between this format and that of the Baldauf studies becomes more apparent.

#### Pilot Study

A pilot study conducted by this writer using a small group of sixth grade students yielded significant reliability scores using the Kuder-Richardson-20 (KR-20) formula on a matching lexical cloze procedure and an informal reading inventory. There was also no significant

difference in scores between groups taking the matching lexical cloze test and informal reading inventory over the same passage and groups having an alternate matching lexical cloze and the same informal reading inventory as the experimental group. Thus, no order effects were indicated.

#### Needed Research

Prior research on the cloze technique has compared it with more traditional or standardized methods. Little research has been conducted to show the validity of using a matching cloze test with students in the intermediate grade levels.

#### Statement of the Problem

The specific purposes of this research are:

- a) to determine the relationship between the matching lexical cloze procedure, two group informal reading inventory procedures, and a standardized reading test (i.e., Stanford Achievement Test, Reading Comprehension),
- b) to determine if the presented matching lexical cloze procedure will produce reliable and valid results, and
- c) to determine when the matching lexical cloze technique can be used by the classroom teacher.

#### Other Purposes

Other purposes of this research are: a) to determine comparable instructional, independent, and

frustration score ranges for the matching lexical cloze procedure in relation to traditional group informal reading inventories, and b) to compare the use of a matching lexical cloze test and the teacher ratings of student ability by placing them in instructional, independent, and frustration levels for specific reading passages.

### Definition of Terms

#### Multiple-Choice Cloze

The term multiple-choice cloze refers to the procedure developed by O'Reilly et al. This procedure contains only lexical deletions and utilizes grammatically correct but semantically implausible distractors selected from word lists. The word lists used included the EDL Research and Information Bulletin 5, A Revised Core Vocabulary and the American Heritage Word Frequency Book. Raw scores on the multiple-choice cloze tests are converted to percentage scores on the test passages developed according to the procedure mentioned above.<sup>11</sup>

#### Matching Lexical Cloze

This term refers to the procedure developed by this writer. This procedure contains only lexical deletions (nouns, verbs, verbals) and utilizes grammatically correct but semantically implausible distractors selected from material immediately preceding the passage.

A deletion rate of approximately every fifth word was attempted. For a complete method of deletion and distractor selection, see Appendix A. Raw scores on the matching lexical cloze tests were converted to percentage scores.

#### Traditional Tests

This term refers to standardized multiple-choice tests administered to students in which they are asked to answer questions concerning material in the passages read. Both literal and inferential questions are included. The test used in this research was the Stanford Achievement Test, Advanced Form B. The raw scores on the Reading Comprehension Subtest were correlated with percentage scores on the matching lexical cloze tests and the group informal reading inventories.

#### Informal Reading Inventory

This term relates to tests constructed which test the student's literal comprehension level. Students are required to read passages of 300 to 500 words and supply short answers to twenty to twenty-five questions based on recall without referring back to the passages.

Finn developed a type of informal reading inventory in 1973 in which objectively formulated questions based on exact passage wording are intended to measure students' comprehension of the passage material.<sup>12</sup> Raw scores are converted to percentage scores

to determine a student's level of comprehension.

#### Wh-Questions

Questions were formulated on Bormuth's theory using exact wording from the test passage but replace the desired response with who, what, where, when, which, and why.<sup>13</sup> The raw score of total correct responses is converted to a percentage score in order to determine a student's level of comprehension.

#### Research Hypotheses

The specific problem to be investigated using eighth grade middle school students is how the matching lexical cloze, two types of informal reading inventories, and a traditional, standardized reading technique compare to each other in terms of difficulty, validity, and reliability as measures of reading comprehension.

The research hypotheses to be examined are that the scores on 1) matching lexical cloze test scores will show a significant positive correlation with scores on a group informal reading inventory, 2) matching lexical cloze test scores will show a significant positive correlation with standardized reading test scores, and 3) a teacher ranking of students' reading ability will show a significant positive correlation with student ranking on the matching lexical cloze tests.

#### Limitations

1. Due to the limited samples examined, gener-

alization to all eighth grade students or any larger population is not possible.

2. Variables such as sex, IQ, and prior method of reading instruction which could affect results were not examined.

3. The matching lexical cloze procedure might have been unfamiliar to most of the students examined.

4. Selection of a different population might yield different results.

5. Both matching lexical cloze and informal reading inventory were developed from the same passage.

Despite these limitations, the results of this study should prove useful in determining the effectiveness of the evaluative instruments examined as measures of reading comprehension and achievement in the regular classroom.

### Footnotes

<sup>1</sup>J. T. Guthrie; M. Seifert; N. A. Burnham; and R. I. Caplan, "The Maze Technique to Assess, Monitor Reading Comprehension," The Reading Teacher, 28 (1974).

<sup>2</sup>J. R. Bormuth, "Comparable Cloze and Multiple-choice Comprehension Test Scores," Journal of Reading, 10 (1967).

<sup>3</sup>R. P. O'Reilly; S. J. Schuder; S. J. Kidder; R. Salter; and P. D. Hayford, The Validation and Refinement of Measures of Literal Comprehension in Reading for Use in Policy Research and Classroom Management (Albany, New York: ERIC Document Reproduction, Ed 133 363, 1976).

<sup>4</sup>J. R. Bormuth, "Cloze Tests as Measures of Readability and Comprehension Ability," (Ph. D. dissertation, University of Indiana, 1962).

<sup>5</sup>O'Reilly et al. (1976).

<sup>6</sup>D. J. Rocchio, "The Validity of Four Non-traditional Measures of Reading Comprehension with Community College Students," (Ed. D. dissertation, University of Missouri-St. Louis, 1978), p. 5.

<sup>7</sup>Ibid.

<sup>8</sup>O'Reilly et al. (1976).

<sup>9</sup>Ibid.

<sup>10</sup>R. B. Baldauf, Jr; R. L. T. Dawson; J. Prior; and I. K. Propst, Jr., "Can Matching Cloze Be Used with Secondary ESL Pupils?" Journal of Reading, 23 (1980) and

<sup>11</sup>Rocchio, p. 11.

<sup>12</sup>P. J. Finn, "An Item Writing Algorithm for Written English," (Ph.D. dissertation, University of Chicago, 1973).

<sup>13</sup>Bormuth, "Comparable Cloze," (1967).

## CHAPTER II

### REVIEW OF RELATED LITERATURE

Since this study concerned itself with the comparison of three types of comprehension measures and the validity of these as measures of reading comprehension, such research will be analyzed here. Research related to the application of non-traditional measures of literal comprehension levels, reliability, and test construction will be reviewed. Non-traditional measures not included in the project or not applicable to it will be omitted.

#### Non-Traditional Reading Measures

Several measures will be examined in this section, primarily the cloze procedures, modified cloze procedures, and variations thereof. Various designs will be examined for reliability, validity, and use as readability guides.

#### Cloze Tests

Taylor introduced the cloze method in 1953.

He defines it as:

A method for intercepting a message from a transmitter, mutilating its language pattern by deleting words, and administering it to receivers in such a way that their attempts to make the patterns

whole again potentially yield a measure of their ability to deal with the general meaning and form intended.<sup>1</sup>

Taylor's method of construction of a cloze test consists of deleting words at random and replacing these words with blanks of some standard length, presenting the mutilated passage to some group of subjects, totaling the correct number of responses, and, thus, obtain readability scores.<sup>2</sup> However, the traditional method which Bormuth popularized consists of deleting every fifth word and demands exact substitution on the part of the student.<sup>3</sup>

#### Alternate Test Constructions

Many researchers have investigated the possibilities of various word deletion patterns. Rubenstein and Sterling found that increasing the context beyond ten words between deletions does not seem to increase the accuracy of word prediction; however, the length at which the context contains maximum effectiveness lies between five and ten words.<sup>4</sup>

Aborn et al. investigated the constraint upon words attributable to the number of words between deletions, distribution of deletions, and structure of the passage. One word was omitted from each sentence in a way that yielded three treatments of sentence length, four treatments of position of omission, and six treatments of word class. The results of their study

indicate:

1. Length and distribution of context are independent sources of constraint.
2. Predictability of words is inversely related to the size of class.
3. Context over ten or under five words produces maximum constraint.
4. When frequency or word class is uncontrolled, predictability is similar for all deletion positions except final which is lower.<sup>5</sup>

Fillenbaum obtained cloze scores utilizing varying deletion rates of 1:2, 1:3, 1:4, 1:5, 1:6. The main finding of this investigation was that cloze scores increased moderately with decreasing density of deletions.<sup>6</sup>

#### Lexical Cloze Deletions

Several researchers have studied the use of a lexical cloze procedure arguing that this type of instrument is a more accurate measure of comprehension since it requires more semantic encoding on the part of the reader than standard nth word deletion measures which emphasize a higher proportion of syntactic encoding.<sup>7</sup>

In summary of his earlier research, Rankin elucidates the advantages of a lexical deletion pattern as:

If we consider the fact that lexical deletions may result in more reading related items than 'every nth' deletions and the fact that every 'nth deletions' are more subject to the influence of intelligence, general reading comprehension and language skills than deletions, then it would follow that the use of this deletion system is less likely to result in significant gains than the use of the measurement

technique which emphasizes the measurement of lexical comprehension.<sup>8</sup>

In a study of subjects in grades three through seven conducted by McLeod using alternate forms on seven passages, a deletion pattern of every eighth word was determined but later modified when certain blanks were determined to produce ambiguous responses. Those blanks were replaced so that only "unequivocal" words were deleted; however, no criterion for unequivocality was stated.<sup>9</sup> Since there is no clear definition for method of deletion selection, this study has been criticized for the practice of selecting cloze deletions by a method other than a predetermined mechanical formula.<sup>10</sup>

Louthan's investigation which used passages which had deletion rates of every tenth word for: a) only nouns, b) only verbs, or c) only adjectives indicated no difference in comprehension scores between groups. However, groups in which only prepositions and conjunctions or pronouns and noun determiners were deleted had significantly higher scores on the cloze test.<sup>11</sup> The results of this study indicate that "content" word deletion tests require different skills on the part of the reader than "function" word deletion tests.

Luke investigated four types of deletion tests with a small group of undergraduate college women. The test types consisted of: nouns only, verbs only,

adjectives only, and a combination of three. The results of the study indicate higher mean scores for verb deletions only which were followed by the mean score on the combination test, followed by nouns only, and followed by the lowest mean score on the adjective only test.<sup>12</sup>

In a study involving 128 college upperclassmen and graduate students, Greene compared a standard cloze deletion procedure with a rate of every 12th word with a lexical procedure which deleted only content words, namely: nouns, verbs, adjectives, and adverbs. The content word tests did show higher scores. Although there was no significant difference in the mean scores between the tests, the lexical cloze test scores were more reliable and less variable than the standard deletion cloze scores. However, Greene pointed out that there was a loss of objectivity in selection of items using this rational deletion method.<sup>13</sup> Other researchers have criticized this method claiming that it ceases to measure passage difficulty since the deletions are no longer representative of the population of possible deletions within the sample.<sup>14</sup>

In a study conducted by Coleman, words were classified in two categories, content words and function words. A standard nth deletion at an approximate twenty percent rate was used as well as varying the initial word deleted. Coleman obtained cloze scores for each word

as well as each word category. He determined that a passage was easier to comprehend when function word deletions appeared more frequently than content word deletions. It was concluded that content words were more difficult to supply because they supposedly carry more information than function words.<sup>15</sup>

#### Multiple-Choice Cloze

This type of test retains the advantages of standard cloze tests, namely objectivity and passage dependency. Rankin claims this type of test also decreases the level of student anxiety involved in a cloze procedure. Multiple-choice cloze tests have more variety of format and scoring than the standard cloze procedures.<sup>16</sup> In some multiple-choice tests synonyms are scored as correct responses. Some researchers believe this method of scoring increases the face validity of the multiple-choice cloze format because it decreases the importance of prior knowledge of the subject as a factor in completing the items.<sup>17</sup>

Other multiple-choice cloze procedures differ in a) type of deletion (lexical or standard nth), b) types of distractors (semantic or syntactic), c) procedure for selecting distractors (word lists, computer banks, random choice, or attached passages), and d) number of distractors (three to five). In this type of procedure, alternate possibilities are offered as well as the correct response for each deleted word.

The advantages of using a multiple-choice cloze format are the construct validity and high reliability particularly in test formats similar to O'Reilly and Streeter's. However, the disadvantage is the lack of exact replicability by two or more individual test developers. It seems highly unlikely that two individuals would choose exactly the same distractors or develop identical test items even if they used identical word lists.

In a study using college sophomores, Cranney used four alternative, multiple-choice cloze tests and a standard deletion rate of every tenth word. Distractors were selected from incorrect responses of students who took a regular cloze test. It should be noted that the distractors were primarily synonyms for the correct responses. Cranney's test showed a moderate correlation with a standardized reading test.<sup>18</sup> This use of semantically plausible distractors has been criticized by researchers who indicate that this method, "produces test items, one-third of which are almost impossible to answer correctly".<sup>19</sup>

A three alternative multiple-choice cloze technique, called a maze technique, was developed by Guthrie et al. The distractors selected consist of one syntactically plausible distractor and one distractor that is both syntactically and semantically implausible. The distractors were selected from the passage itself

and word lists. Four categories of words a) nouns and pronouns, b) verbs and auxiliaries, c) adjectives and adverbs, and d) prepositions, conjunctions, and articles were deleted. Results indicated a high reliability on all seven passages.<sup>20</sup>

A maze test developed at the fifth to sixth grade level designed for administration to fifth graders was designed by Pikulski and Pikulski using a standard fifth word deletion rate. Maze test results and traditional cloze test results were compared to teacher evaluation and reading group placement with respect to students' independent, instructional, and frustration levels. Both the maze and the traditional cloze tests were found to over-estimate student placement in reading groups when compared to teacher judgements. However, the cloze test correctly placed sixty-seven per cent of the students in reading groups; whereas, the maze correctly placed fifty per cent of the students.<sup>21</sup>

A lexical deletion pattern of nouns, verbs, adjectives, and adverbs was investigated by O'Reilly et al. This design was selected due to the fact that these word classes carry most of the information in a passage and, therefore, an increase of semantic comprehension would be measured with a reduction in the syntactic comprehension measurement. This study utilized distractors which are syntactically acceptable but semantically unacceptable. In this manner, the

reader must know that the distractors do not fit the meaning of the sentence context or the surrounding context.<sup>22</sup>

In factor analysis studies of this format, O'Reilly and Streeter found it a valid measure of literal comprehension that required both semantic and syntactic analysis and comprehension.<sup>23</sup>

#### Matching Cloze

In a study using secondary ESL students, Baldauf et al. found a matching cloze procedure produced tests which were reliable (.88 using KR-21) and valid measures of reading comprehension. The tests consisted of passages of 300 to 400 words divided into ten sections of approximately 30 to 40 words each. Approximately every fifth word was deleted, and the deletions were determined to be mutually exclusive. The five correct options were listed in the margin for each section. No distractors were used in these tests.<sup>24</sup> This type of test can be easily constructed by the classroom teacher since the selection of distractors is not necessary.

In a similar study using elementary ESL students, Baldauf and Propst used both matching and multiple-choice cloze tests. Correlations ranging from .5 to .8 were found between the matching cloze and the multiple-choice cloze tests used and the Gates MacGinitie Vocabulary and Comprehension Test. The findings of this

study indicate a higher internal consistency of the matching cloze tests over the multiple-choice cloze tests.<sup>25</sup>

### Comparison of Cloze Formats and Standardized Tests

#### Validity

Many cloze technique researchers have compared their test results with scores on standardized measures of reading comprehension or intelligence. These comparisons were often used to establish concurrent validity; however, certain questions remain concerning the actual validity of some standardized measures as tests of literal level comprehension.

The Stanford Achievement Paragraph Meaning Test was compared to a cloze technique by Ruddell using elementary students. The corresponding correlations ranged from .61 to .74.<sup>26</sup>

Jenkinson compared high school students' scores with the Cooperative Reading C2, Vocabulary and Comprehension Test with correlations of .78 and .73 respectively.<sup>27</sup>

In other studies Friedman reported correlations of .71 to .87 between cloze procedures and the Metropolitan Achievement Test,<sup>28</sup> and Schneyer achieved correlations of .60 to .68 with cloze procedures and the Gates Reading Survey.<sup>29</sup>

Hafner conducted a study utilizing different

methods of scoring: a) connective words only, b) content words, c) content/connective quotient, d) grammatically correct but lexically incorrect, and e) incongruous responses. The cloze scores correlated positively and significantly with the Michigan Vocabulary Profile; the Otis Quick Scoring Mental Ability Test, Gamma Form; and the Weschler-Bellevue Intelligence Scale, Information.<sup>30</sup>

Findings of O'Reilly and Streeter indicate that the multiple-choice cloze procedure:

is in part a measure of a restricted form of reading comprehension that is essentially independent of IQ. This form of comprehension appears to be interpretable as the apprehension of the "strictly literal" meanings contained in sentences and phrases as measured by reading tests that focus on factual questions, questions about explicit details, and questions about interpretation of meanings within the context.<sup>31</sup>

This study compared the multiple-choice cloze format scores with informal reading inventory scores, the Short Form Test of Academic Aptitude, and the California Achievement Test scores. The sample used consisted of students in grades one through six in a medium-sized urban school district.<sup>32</sup>

Smith and Zinc compared cloze tests with the subtests of the Davis Reading Test, Form 2A (DRT) and concluded that the cloze test:

required a student to draw upon abilities in literal comprehension, comprehension of main ideas, inferential comprehension, and structural awareness to a degree moderately similar to the application of subskills required by the multiple-choice format of the DRT.... The findings provide

evidence of the validity of the cloze test as a measure of reading comprehension and its subskills as measured by a conventional standardized reading achievement test.<sup>33</sup>

### Cloze Tests as Measures of Readability

Experts and investigators alike accept Betts' model for reading comprehension levels (informal reading inventories) of ninety percent as the independent level, seventy-five percent as the instructional level, and fifty percent as the frustration level.<sup>34</sup>

Bormuth conducted a study in 1968 which determined that cloze scores of forty-four percent and fifty-seven percent corresponded to Betts' seventy-five percent and ninety percent levels on a standard every fifth word deletion rate cloze test. Bormuth concluded that:

a student can gain very little information from studying materials on which his cloze readability score is below thirty-seven percent and that using materials much easier than the thirty-seven percent level will permit the child only slightly more knowledge while reading them.<sup>35</sup>

In a replication of Bormuth's study concerning readability levels, Rankin and Culhane found corresponding cloze levels of forty-one percent and sixty-one percent met Betts' criteria for instructional and independent levels of comprehension.<sup>36</sup>

Heaney found cloze tests to be ninety-two percent reliable as determinants of readability. She criticizes the more traditional measures or formulas in that they fail to consider:

the following important factors involved in passage comprehension: a) the impact of simple structure words on understanding the interrelationship of ideas, b) multiple meanings of simple key vocabulary terms as attributes of a concept, c) idea density (abstractions expressed in simple words), and d) the specialized patterns of writing inherent to a discipline.<sup>37</sup>

### Objectivity of Cloze Formats and Informal Reading Inventories

Most researchers agree that the cloze or modified cloze format is generally an objective method of test item construction, the standard nth word deletion pattern being the most objective method, and the word list consultation or item developer's judgment being the least objective. However, these methods are still far less subjective than the traditional method of teacher-developed questions commonly used with classroom materials and informal reading inventories.

In an attempt to remove this subjectivity of question formulation, Bormuth designed test questions which are termed wh-questions. With this type of question formulation, the interrogative pronouns "who", "what", or "by whom" are inserted in the grammatical constructions found in the text or passage itself. An example of this method follows:

The boy rode the steed. (original sentence)

Who rode the steed? (wh-question)

Who rode the horse? (semantic wh-transform)

This type of question construction is far more

objective in nature as a test of literal comprehension than previously designed item construction.<sup>38</sup>

Finn has developed a variation of Bormuth's wh-question procedure which is termed Finn's Algorithm.

Finn defines reading achievement as:

....the ability to answer questions that measure information gain (i.e., reading achievement) is evidence that a person has comprehended and been able to retrieve in the form demanded by the question an item of information encoded in the written text from which the question was generated. In short, achievement is evidence of comprehension plus retrieval....Failure on questions is evidence of failure of achievement and perhaps--but not necessarily--a failure of comprehension.<sup>39</sup>

Finn's criteria for test development include the following:

1. The questions are generated by rules such that no undirected decisions are permitted on the part of the item writer. If item writers working independently produce the same questions from the same texts, the objective will be accomplished.
2. What is an acceptable answer, is rule-generated and derived from the rules governing the question.
3. The definitions must be such that one can persuasively argue that success or failure on answering questions is a useful definition of reading achievement, and is therefore related to information gained.
4. The percentage of items answered correctly would reflect the percent of information mastered in the passage. The implication is that there can be a one to one relationship between questions and items of information.
5. The procedure for selecting questions and therefore units of information would insure that the test measures a representative sample of the information in the passage.<sup>40</sup>

The claim that this algorithm produces questions that can be replicated by various item writers was substantiated in Finn's study when seventy-seven times out of one-hundred item writers produced questions with

exactly the same words in exactly the same order.<sup>41</sup>

Finn also allowed synonyms as acceptable responses. He does, however, acknowledge:

the fact that the reader may be able to answer some items generated...without any understanding but states that the likelihood of this phenomenon is small when given lengthy passages (i.e., 300 words) and free recall questions.<sup>42</sup>

In addition, Finn did not test his assumption that his questions were passage dependent. It may be possible for some readers to correctly respond to some passage questions without actually having read or comprehended the material contained in those passages. However, Finn did employ a rational rather than random method for identifying key information units.<sup>43</sup>

This method of item generation seems to be an objective, operational, and rational method for determining information gain. It measures:

...in gross terms a gain of information stated explicitly within clauses, with the exception of references to or dependence upon other clauses that can be specified in grammatical terms.<sup>44</sup>

Researchers have raised questions concerning whether this type of item measures literal understanding or short term memory.<sup>45</sup> However, this has not been empirically determined.

Most questions regarding the objectivity of questions and students' ability to respond to them is summarized by Guthrie et al. This study questions the ability of the subject to correctly answer questions as follows:

If a child fails to answer one or more questions, it is difficult to determine whether he did not understand the question, did not understand the passage, failed to relate the question to the proper section of the passage, or some combination of these... Is the child's incorrect response due to inconsistency in the difficulty of the questions rather than increases or decreases in his understanding of the material...or does the response made rather than the conceptual complexity determine the difficulty of the question.<sup>46</sup>

A final question raised by many researchers is can the question be answered without prior reading of the passage.

#### Order Effects

Bormuth's study in 1967 revealed no order effects when students were administered a standard cloze test followed three days later with an informal reading inventory constructed from the same passage. Bormuth found no significant difference in mean scores between those students with double exposure to identical passages (i.e., cloze test followed by informal reading inventory) and the mean scores of students who had not had prior exposure to the informal test through a cloze task.<sup>47</sup>

#### Summary

Although several deletion methods may be employed, the most often used is the nth method, and the most frequent is an every fifth word deletion pattern. While this procedure appears to be an adequate measure for determining passage difficulty or

readability, the best method for determining reading comprehension levels for individual students has yet to be established. Other deletion methods include ten percent to twenty percent random deletion patterns, deletions based on form class, and deletions based on the rational selection of the investigator to be critical to the meaning transmitted.

Most often the formula used to determine passage difficulty was the Dale-Chall.<sup>48</sup> In many studies the publisher's estimate of passage difficulty was used.

Other instruments used to validate cloze results included standardized achievement or intelligence tests and, in several studies, specially constructed multiple-choice items.

Several investigators used scoring methods other than exact word substitutions including synonyms, form class similarities, and grammaticality. The comparison of cloze scores with standardized scores seems to hold up well and was reported as significant in nearly all studies. In addition, all studies have shown that there seems to be a positive relationship between a subject's ability to use context clues, either semantic or syntactic, and his ability to comprehend written or oral discourse. However, deletions more closely spaced than five words tend to greatly increase the difficulty of restoring the exact

word deleted, and cloze scores tend to increase moderately with a decreasing density of deletions.<sup>49</sup>

A scoring technique which accepts only exact word substitutions seems to correlate most significantly with standardized measures. Also, it was found that word patterns of material written in the active verb form were more predictable and easier for test subjects to comprehend, thus, yielding higher cloze scores.<sup>50</sup>

### Conclusions

Research suggests that the most reliable and valid cloze test scores are obtained with the following:

1. A mechanical mutilation system is followed.
2. Not more than twenty words out of every one hundred are deleted.
3. Longer passages yield better results.
4. Deletion rates of every tenth to twelfth word in longer passages may be valid for certain purposes.
5. Scoring exact word substitutions for deleted words is the most efficient scoring criteria.
6. Scoring of form class or content and function words may provide more specialized results and information.

### Pilot Study

A pilot study with sixth grade students in which a matching lexical cloze procedure was utilized that

deleted only nouns and verbs was conducted by this writer. This matching lexical cloze test was followed by a group informal reading inventory over the same passage. Reliabilities based on the Kuder-Richardson (KR-20) formula were .71 on the matching lexical cloze procedure and .69 on the informal reading inventory. Both test passages were constructed from the "Coal II" passage of Finn's study.

A relatively high correlation with the cloze test and the Stanford Achievement Test, Intermediate Level, Form B, .72, was determined. However, correlations between the cloze scores and the informal reading inventory were only moderate, (.45).

The results of this study indicate further research is needed in an attempt to estimate readability levels, reliability, and validity for this type of procedure.

J. J. Kessler, Department of Psychology,  
University of Mississippi, Oxford, Mississippi

J. Method and J. Assessment, "Reading-  
Assessment and Self-Regulation of Learning,"  
Journal of Educational Psychology, 54 (1962), 2-11.

Potter, J.

1971.

J. Loring, "Word Recognition and Reading-  
Comprehension and Their Effects on Reading-  
Comprehension," Journal of Educational Psychology, 54 (1962), 12-21.

Footnotes

<sup>1</sup>T. L. Taylor, "Cloze Procedure: A New Tool for Measuring Readability," Journalism Quarterly, 30 (1953), p. 415-433.

<sup>2</sup>T. C. Potter, "A Taxonomy of Cloze Research, Part I: Readability and Reading Comprehension," Washington, D.C.: ERIC Document Reproduction, Ed 035 514, (1976), p. 4.

<sup>3</sup>J. R. Bormuth, "Cloze Tests as Measures of Readability and Comprehension Ability," (Ph. D. dissertation, University of Indiana, 1962).

<sup>4</sup>Potter, p. 4.

<sup>5</sup>M. Aborn; H. Rubenstein; and T. D. Sterling, "Sources of Contextual Constraint upon Words in Sentences," Journal of Experimental Psychology, no. 3 (1959): p. 117.

<sup>6</sup>S. Fillenbaum; L. V. Jones; and A. Rapoport, "The Predictability of Words and their Grammatical Classes as a Function of Rate of Deletion from a Speech Transcript," Journal of Verbal Learning and Verbal Behavior, 2 (1963): 186-194.

<sup>7</sup>D. J. Rocchio, "The Validity of Four Non-traditional Measures of Reading Comprehension with Community College Students," (Ed.D. dissertation, University of Missouri-St. Louis, 1979), p. 56.

<sup>8</sup>J. McLeod and J. Anderson, "Readability Assessment and Word Redundancy of Printed English," Psychological Reports 18 (1966): 35-38.

<sup>9</sup>Potter, p. 10.

<sup>10</sup>Ibid.

<sup>11</sup>V. Louthan, "Some Systematic Grammatical Deletions and their Effects on Reading Comprehension," English Journal, 54 (1965): 295.

<sup>12</sup>"Form Class and Cloze Procedure," quoted in Potter, p. 17.

<sup>13</sup>F. P. Greene, "Modification of the Cloze Procedure and Changes in Reading Test Performances," Journal of Educational Measurement 2 (1965): 213-217.

<sup>14</sup>Potter, p. 26.

<sup>15</sup>"Cloze Scores of Nominalizations and their Grammatical Transformations Using Active Verbs," quoted in Rocchio, p. 58.

<sup>16</sup>E. F. Rankin, "The Cloze Procedure Revisited," in *Interaction: Research and Practice for College-adult Reading*, Twenty-third Yearbook of the National Reading Conference, P. L. Nacke (Ed.) National Reading Conference, Inc., (1974).

<sup>17</sup>Rocchio, p. 61.

<sup>18</sup>A. G. Cranney, "The Construction of Two Types of Cloze Reading Tests for College Students," Journal of Reading Behavior 51 (1972-73): 60-64.

<sup>19</sup>Rocchio, p. 62.

<sup>20</sup>J. T. Guthrie; M Seifert; N. A. Burnham; and R. I. Caplan, "The Maze Technique to Assess, Monitor Reading Comprehension," The Reading Teacher 28 (1974): 161-168.

<sup>21</sup>J. J. Pikulski and E. C. Pikulski, "Cloze, Maze and Teacher Judgment," The Reading Teacher 30 (1977): 766-770.

<sup>22</sup>The Validation and Refinement of Measures of Literal Comprehension in Reading for Use in Policy Research and Classroom Management, quoted in Rocchio, p. 64.

<sup>23</sup>R. P. O'Reilly and R. C. Streeter, "Report on the Development and Validation of a System for Measuring Literal Comprehension in a Multiple-choice Cloze Format: Preliminary Factor Analytic Results," Journal of Reading Behavior 9 (1977).

<sup>24</sup>R. B. Baldauf, Jr; R. L. T. Dawson; J. Prior; and I. K. Propst, Jr., "Can Matching Cloze Be Used with Secondary ESL Pupils?," Journal of Reading, 23 (1980) 435-440.

<sup>25</sup>R. B. Baldauf, Jr. and I. K. Propst, Jr., "Matching and Multiple-choice Cloze Tests," Journal of Educational Research, vol. 72, no. 6 (July-August 1979): 321-136.

<sup>26</sup>"A Study of the Cloze Comprehension Technique in Relation to Structurally Controlled Reading Material," quoted in Potter, p. 6.

<sup>27</sup>M. E. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension," (Ph.D. dissertation, University of Chicago, 1957).

<sup>28</sup>M. B. Jones and E. C. Pikulski, "Cloze for the Classroom," Journal of Reading 17 (1974): 432-438.

<sup>29</sup>J. W. Schneyer, "Use of the Cloze Procedure for Improving Reading Comprehension," Reading Teacher 19, (1965): 174-179.

<sup>30</sup>L. E. Hafner, "Cloze Procedure," Journal of Reading 9 (1966): 415-421.

<sup>31</sup>O'Reilly and Streeter, p. 67.

<sup>32</sup>Ibid.

<sup>33</sup>N. Smith and A. Zinc, "A Cloze-based Investigation of Reading Comprehension as a Composite of Subskills," Journal of Reading Behavior 9.4 (1977): 398.

<sup>34</sup>E. A. Betts, Foundations of Reading Instruction, New York: American Book Co., 1946.

<sup>35</sup>J. R. Bormuth, "Cloze Readability Procedure," Elementary English, 45 (1968): 433.

<sup>36</sup>E. F. Rankin and J. W. Culhane, "Comparable Cloze and Multiple-choice Comprehension Test Scores," Journal of Reading 13 (1969-70): 193-198.

<sup>37</sup>L. Heaney, The Cloze Procedure, Fond du Lac, Wisc.,: ERIC Document Reproduction, Ed 151 762, (1978): 48.

<sup>38</sup>J. R. Bormuth, On the Theory of Achievement Test Items, Chicago: University of Chicago Press, 1970.

<sup>39</sup>P. J. Finn, "An Item Writing Algorithm for Written English," (Ph. D. dissertation, University of Chicago, 1973).

- <sup>40</sup>Bormuth, "Comparable Cloze," (1967).
- <sup>41</sup>Finn, p. 147-148.
- <sup>42</sup>Rocchio, p. 33
- <sup>43</sup>Ibid., p. 34.
- <sup>44</sup>Finn, p. 38.
- <sup>45</sup>Rocchio, p. 35.
- <sup>46</sup>Guthrie et al., p. 163.
- <sup>47</sup>Bormuth, "Comparable Cloze," (1967).
- <sup>48</sup>E. Dale and J. Chall, "A Formula for Predicting Readability," Educational Research Bulletin 27 (1948): 11-28.
- <sup>49</sup>Potter, p. 38.
- <sup>50</sup>Ibid.

### CHAPTER III

#### METHODS AND PROCEDURES

The factors affecting this research design include the selection of a population, the selection and adaptation of suitable materials, the development of various test forms, and the acquisition of standardized test data on all subjects.

##### Sample Used

The subjects comprising the sample for this research included thirty-five students (18 males/17 females) enrolled in eighth grade language arts classes in the Lincoln County R-IV School District's Middle School. The average group IQ score was 106. The school district is located in a primarily rural area north of St. Charles County. Ninety-one percent of the working population of the district are blue collar workers with nine percent being classified as professional workers. Eighty-two percent of the population commutes to the metropolitan St. Louis area for employment. The average income of families within the ninety-seven square mile district is at the state average of \$12,500. The Middle School has an enrollment

of 255 students in grades six through eight. The total enrollment for the school district is 1,046.

### Instruments and Materials Used

#### Matching Lexical Cloze

Two matching lexical cloze tests at the sixth and eighth grade levels with forty deletions were administered on consecutive days. A brief practice session preceded the initial matching lexical cloze test.

The tests were constructed using selections from Scott-Foresman Basics in Reading Series and passages from Finn's study.<sup>1</sup> The readability formula used to estimate passage difficulty was the Fry formula.<sup>2</sup> The cloze passage from the reading texts totalled 309 words, and the passage from the Finn study was 349 words.

The matching lexical cloze tests included material concerning "tigers" and "coal" in the sixth grade passages. "One man's initiative" and "pirates" were the topics of the eighth grade passages. These tests consisted of two pages of lexically clozed items (nouns, verbs, and verbals) with twenty deletions per page. The twenty correct responses as well as six distractors were listed in random order at the bottom of each page. The complete method for deletion and selection of distractors is found in the Appendix. Distractors were selected in a reverse cloze method

beginning with the sentence immediately preceding the the clozed passage.

In each of the matching lexical cloze tests, the passages began with two complete sentences with no deletions. The deletions began in the third sentence of the passage. All passages began with the initial sentence of a paragraph. The passages were clozed for nouns, verbs, and verbals only. Words were deleted as nearly as possible to a 1:5 deletion ratio. (see Appendix) All clozed words were listed in random order at the bottom of the page.

Distractors were selected in a reverse cloze method beginning with the sentence immediately preceding the passage. If necessary, distractors were altered in number or tense only to agree with deleted words.

#### Finn's Informal

One group informal reading inventory was constructed from the Finn passages "Coal II" and "Pirates". These passages were 514 and 473 words in length respectively and were identical to the passages and questions used by Finn.

Finn generated twenty-five wh-questions for the "Coal II" passage and twenty wh-questions regarding the "Pirate" passage.<sup>3</sup>

This informal reading inventory was administered to the sample group approximately two weeks after the

matching cloze tests. Previous research has shown no order effects resulting from this method of test administration.<sup>4</sup> A Fry readability formula produced a sixth grade level for the "Coal II" passage and an eighth grade level on the "Pirate" passage.

#### Bormuth's Wh-Questions

Another group informal reading inventory was administered one week after the Finn group informal reading inventory. This test was constructed from the matching lexical cloze passages from the Scott-Foresman texts. The method of item construction was Bormuth's wh-method in which interrogative words are substituted in exact sentences from the text, and students are required to supply the necessary information.<sup>5</sup> Twenty questions for each passage were developed in this matter. The passages used consisted of 364 words in the sixth grade passage and 348 words in the eighth grade passage.

#### Stanford Achievement Test

Between the administration of the matching cloze tests and the group informal reading inventories, students were administered the Stanford Achievement Test, Advanced Battery, Form B as part of the school district's testing policy. This test consists of nine tests administered on three separate days. These tests consisted of multiple-choice items. The test is a group-administered, timed, standardized test.

The specific subtest of this test battery that was used in this research was the Reading Comprehension section which consisted of 74 items. According to the Norms Booklet for the test battery, the reliability for this subtest is .94 using the KR-20 formula and a construct validity of .95 is reported. This subtest correlates highly with the other subtests of the battery with a range of .70 to .97.

#### Procedures

All students were tested in two groups in the same classroom during the Spring of 1980. The classroom is average size and well-lighted. It was their regular language arts classroom. Students were seated in assigned class seats. A brief sample exercise for the matching cloze procedure at a fourth grade level of readability was administered and discussed which took approximately ten to fifteen minutes. The instructor walked around the room to see that everyone knew how to complete the task. The initial matching lexical cloze passage from the Scott-Foresman texts was then administered. Following the oral reading of directions by the instructor, students had the remaining forty-five minutes of the class to complete the cloze test.

The following day the second matching lexical cloze test from the Finn material was administered. No practice exercises were offered for this test

session; however, the instructor reminded the students of the previous day's activities. The instructor read the directions for the test and the students were instructed to begin.

On both matching lexical cloze tests separate answer sheets which were numbered consecutively from one to forty to correspond to the deleted blanks were used by students to mark responses.

The following week students were administered the Stanford Achievement Test, Advanced Battery, Form B in three testing sessions. Standardized instructions were followed in the administration of the test.

The following week a group informal reading inventory from the Finn study was administered to the sample group. Students were instructed to read the passage carefully and to be ready to respond to several questions concerning the material contained in the passages following the reading. Students were informed that they would not be allowed to look back to the passages for answers, but they could read the passage twice if they wished. Answer sheets numbered from one to twenty-five and one to twenty were distributed for the recording of responses prior to test distribution. Students were instructed to raise their hands when reading was completed. The instructor then collected the passage and handed out the appropriate questions to each student. This procedure was followed for both

sections of the test. One week later the group informal reading inventory developed from the Scott-Foresman material was administered following the same procedure.

#### Data Analysis

Total raw scores for all tests except the Stanford Achievement Test were converted to percentage scores. The reliability scores for the tests taken by the sample group using the KR-20 formula produced results of .98 for all four tests using the matching lexical cloze procedure and the group informal reading inventories. Correlations between the matching lexical cloze scores and the group informal reading inventory scores were obtained using the Pearson Product-Moment correlation coefficient formula. These correlations were only moderate (.49 and .50); however, a correlation of .76 was found between the two group informal reading inventories.

Based on Bormuth's method of determining independent, instructional, and frustration levels for the conventional cloze procedures<sup>6</sup>, it was determined that similar levels for the matching lexical cloze design used in this study could be determined utilizing a visual inspection/averaging technique.

In addition, teacher ranking of students was compared to actual test performance as well as a teacher rating of students' reading ability levels. The correlations

for these ranged from .60 to .74 and .46 to .54 respectively.

Cloze tests were scored by accepting only the exact responses deleted or minor misspellings of these. Synonyms were scored as correct on the group informal reading inventories as well as exact responses from the passages. Misspellings were also scored as correct provided the answer the student intended was apparent.

Footnotes

<sup>1</sup>P. J. Finn, "An Item Writing Algorithm for Written English" (Ph. D. dissertation, University of Chicago, 1973).

<sup>2</sup>E. Fry, "A Readability Formula that Saves Time," Journal of Reading 11 (1967).

<sup>3</sup>Finn, (1973).

<sup>4</sup>J. R. Bormuth, "Comparable Cloze and Multiple-choice Comprehension Test Scores," Journal of Reading 10 (1967).

<sup>5</sup>J. R. Bormuth, On the Theory of Achievement Test Items, Chicago: University of Chicago Press, 1970.

<sup>6</sup>Bormuth, "Comparable Cloze" (1967).

## CHAPTER IV

### ANALYSIS OF DATA AND FINDINGS

Results obtained will be presented for the three research hypotheses as well as the two additional purposes of this research.

#### Research Hypothesis One

The matching lexical cloze test will show a significant positive correlation with both group informal reading inventory procedures.

#### Statistical Hypothesis One

There will not be a significant positive correlation between the matching lexical cloze tests and the group informal reading inventory procedures.

A Pearson product-moment correlation indicated that the matching lexical cloze test and the group informal reading inventory based on the Finn passages correlated at the .44 level,  $p < .05$ . The alternate form tests constructed from the Scott-Foresman texts correlated at the .49 level,  $p < .05$ .

The correlation of the Finn matching lexical cloze test and the text informal reading inventory was .50,  $p < .05$ , as was the correlation between the

matching lexical test from the text and the informal reading inventory constructed from the Finn material (see table 1).

This shows a significant degree of correlation between these measures. On the basis of these results, the statistical hypothesis of no significant correlation is rejected and the research hypothesis is accepted at the  $p \leq .05$  level.

TABLE 1

CORRELATIONS OF TOTAL SCORES  
ON NON-TRADITIONAL TESTS

	F-IBI <sup>#</sup>	T-IRI <sup>##</sup>
F-MLC <sup>+</sup>	.44*	.50*
T-MLC <sup>++</sup>	.50*	.49*

<sup>#</sup>F-IRI is the group informal reading inventory constructed from the Finn material.

<sup>##</sup>T-IRI is the group informal reading inventory constructed from the Scott-Foresman texts.

<sup>+</sup>F-MLC is the matching lexical cloze test constructed from the Finn material.

<sup>++</sup>T-MLC is the matching lexical cloze test constructed from the Scott-Foresman texts.

\* $p \leq .05$ .

Research Hypothesis Two

The matching lexical cloze test will show a significant positive correlation with a standardized reading test.

### Statistical Hypothesis Two

There will not be a significant positive correlation between the matching lexical cloze test and a standardized reading test.

The Finn matching lexical cloze test showed a .51 correlation with the Stanford Achievement Test, Advanced Battery, Form B, Reading Comprehension subtest using the Pearson product-moment correlation coefficient formula, and the matching lexical cloze test constructed from the Scott Foresman texts showed a correlation of .60,  $p < .05$ .

The group informal reading inventories correlated with the standardized test at only slightly higher levels,  $r = .52$  for the Finn and  $r = .60$  for the test constructed from the texts.

Since these correlation coefficients show a significant correlation at the  $p \leq .05$  level, research hypothesis two of significant correlation is accepted, and statistical hypothesis two is rejected (see table 2).

### Research Hypothesis Three

A teacher ranking of reading ability will show a significant positive correlation with the ranking on the matching lexical cloze tests.

### Statistical Hypothesis Three

A teacher ranking of reading ability of students will not show a significant positive correlation with

the ranking on the matching lexical cloze tests.

TABLE 2  
CORRELATIONS OF TOTAL SCORES ON  
NON-TRADITIONAL TESTS AND THE  
STANFORD ACHIEVEMENT READING  
COMPREHENSION SUBTEST

	SAT#
F-MLC	.51*
T-MLC	.60*
F-IRI	.52*
T-IRI	.60*

#SAT is the Stanford Achievement Test, Advanced Battery, Form B, Reading Comprehension Subtest.

\*  $p \leq .05$ .

Correlation coefficients were calculated utilizing Spearman's Rho formula and yielded correlations of .47,  $p \leq .05$ , on the sixth grade Finn matching lexical cloze test and .59,  $p \leq .05$ , on the eighth grade passage. Correlations on the tests devised from the Scott-Foresman texts were .46,  $p \leq .05$ , on the sixth grade passage and .54,  $p \leq .05$ , on the eighth grade passage.

Given these results, research hypothesis three regarding significant positive correlation is accepted and statistical hypothesis three is rejected (see table 3).

TABLE 3  
 RANKED CORRELATIONS ON THE  
 MATCHING LEXICAL CLOZE  
 AND TEACHER RANKINGS

	RHO Correlations
F-MLC 6 <sup>#</sup>	.47*
F-MLC 8 <sup>##</sup>	.59*
T-MLC 6 <sup>+</sup>	.46*
T-MLC 8 <sup>++</sup>	.54*

<sup>#</sup>F-MLC 6 is the sixth grade passage using the Finn material.

<sup>##</sup>F-MLC 8 is the eighth grade passage using the Finn material.

<sup>+</sup>T-MLC 6 is the sixth grade passage devised from the Scott-Foresman texts.

<sup>++</sup>T-MLC 8 is the eighth grade passage devised from the Scott-Foresman tests.

\* $p \leq .05$ .

#### Additional Purposes One

Comparable instructional, independent, and frustration score ranges were determined for the matching lexical cloze tests in relation to traditional group informal reading inventories.

Regressions calculated comparing scores on the group informal reading inventories and the matching lexical cloze tests indicated scores of 75-80 percent on the matching lexical cloze tests from the Finn

passage were equivalent to scores of 90-100 percent on the group informal reading inventory (independent level); 61-73 percent was equivalent to scores of 60-89 percent (instructional level); and scores of 59 percent and below were equivalent to scores of below 60 percent (frustration level). For complete results, see table 4.

The regression scores for the matching lexical cloze tests and the group informal reading inventories developed from the Scott-Foresman texts were somewhat higher: scores of 81-86 percent on the matching lexical cloze indicated the independent level, 67-79 percent indicated the instructional level, and below 66 percent indicated the frustration level (see table 4).

A visual inspection and comparison of average scores on the matching lexical cloze tests and the group informal reading inventories at the same level produced somewhat different results. On the matching lexical cloze test from the Finn material at the sixth grade level, no students scored in the independent range of 90-100 percent. The average score for the instructional range of 60-89 percent was 77 percent on the matching lexical cloze and 69 percent for the frustration level of below 60 percent. On the eighth grade passage, the scores ranged from 59 percent for the independent level, 51 percent for the instructional level, and 43 percent for the frustration level. When total test scores were compared, scores of 65 and 56 percent on the matching

lexical cloze test were the average scores corresponding to the instructional and frustration level scores on the group informal reading inventories (see table 5).

TABLE 4

## REGRESSION RANGES AND SCORES

	F-MIC	T-MIC
90-100		
Independent	75-80	81-86
60-89		
Instructional	61-73	67-79
Below 60		
Frustration	-59	-66
REGRESSION SCORES		
90	75	81
70	66	71
60	61	67
50	57	62

On the passages developed from the Scott-Foresman texts, a visual inspection and comparison of average scores on the matching lexical cloze tests and the group informal reading inventories at the same levels were conducted. On the sixth grade level, scores of 92, 76, and 52 percent were the average scores for the

independent, instructional, and frustration levels. On the eighth grade passage scores of 88, 67, and 62 percent were the averages for the independent, instructional, and frustration levels. When total test score averages were compared, scores of 83, 72 and 63 percent were the average scores for the independent, instructional, and frustration levels (see table 5).

TABLE 5

## VISUAL INSPECTION (AVERAGES)

## MLC AND IRI SCORES

## AT THE SAME LEVELS

	90-100 *	60-89 *	-60 *
F-MLC 6 #		77	69
F-MLC 8 ##	59	51	43
F-MLC T ###		65	56
T-MLC 6 +	92	76	53
T-MLC 8 ++	88	67	62
T-MLC T +++	83	72	63

\*90-100, 60-89, -60 are scores on the group informal reading inventories which correspond to the independent, instructional, and frustration levels.

#F-MLC 6 is the sixth grade matching lexical cloze passage.

##F-MLC 8 is the eighth grade passage.

###F-MLC T is the total score for both passages.

+T-MLC 6 is the sixth grade passage.

++T-MLC 8 is the eighth grade passage.

+++T-MLC T is the total for both passages.

Additional Purposes Two

A comparison of the use of the matching lexical cloze tests and teacher ratings of students' ability with reference to placement of students in independent, instructional, and frustration levels for specific reading passages was made.

On the sixth grade matching lexical cloze tests, the students were correctly placed according to teacher ratings 60 percent of the time for the independent level; 70 percent for the instructional level; and 70 percent for the frustration level (see table 6).

On the eighth grade matching lexical cloze tests, the students were correctly placed according to teacher ratings 17 percent at the independent level; 55 percent at the instructional level; and 65 percent at the frustration level (see table 6).

None of the matching lexical cloze tests overestimated teacher ratings on the independent level for the sixth grade passages, but did overestimate 14 percent on the instructional level and 30 percent on the frustration level for these passages (see table 6).

On the eighth grade passages the matching lexical cloze tests overestimated teacher ratings 6 percent at the independent level and 39 percent at the frustration level (see table 6).

For the sixth grade passages, the matching lexical cloze tests underestimated student placement

according to teacher ratings 40 percent on the independent level and 16 percent on the instructional level but never on the frustration level (see table 6).

On the eighth grade passages, the matching lexical cloze tests underestimated 84 percent on the independent level, 43 percent on the instructional level, but none at the frustration level (see table 6).

TABLE 6  
MLC ACCURACY OF PREDICTION OF READING  
LEVEL (MLC SCORES COMPARED WITH  
TEACHER RATINGS)

	MLC6*	MLC8**
Independent level:		
Correct prediction	60	17
Underestimation	40	84
Overestimation	0	0
Instructional level:		
Correct prediction	70	55
Underestimation	16	43
Overestimation	14	6
Frustration level:		
Correct prediction	70	65
Overestimation	30	39
Underestimation	0	0

\*MLC6 is the sixth grade matching lexical cloze tests.

\*\*MLC8 is the eighth grade matching lexical cloze tests.

Similar calculations using the scores on the group informal reading inventories were completed. These results can be found in table 7. These results indicated that the matching lexical cloze tests apparently are better predictors of the instructional level than group informal reading inventories when teacher judgment of placement is the criterion.

TABLE 7  
IRI ACCURACY OF PREDICTION OF READING  
LEVEL (IRI SCORES COMPARED WITH  
TEACHER RATINGS)

	IRI6*	IRI8**
Independent level:		
Correct prediction	20	67
Underestimation	80	33
Overestimation	0	0
Instructional level:		
Correct prediction	42	53
Underestimation	56	41
Overestimation	2	6
Frustration level:		
Correct prediction	100	87
Overestimation	0	14
Underestimation	0	0

\*IRI 6 is the sixth grade group informal reading inventories.

\*\*IRI8 is the eighth grade group informal reading inventories.

Independent, instructional, and frustration levels were determined from a comparison of group informal reading inventory scores and matching lexical cloze test scores through regressions and visual inspection. It was determined that ranges of 77-100 percent on the matching lexical cloze tests corresponded to Betts' independent level and that scores of 65-76 percent on the matching lexical cloze tests corresponded to Betts' instructional level. Scores of 64 and below on the matching lexical cloze tests corresponded to the frustration level on a group informal reading inventory.

#### Additional Findings

All tests were found to have high reliability ranging from .94-.98 using the KR-20 Formula. These results can be found in table 8.

TABLE 8

#### RELIABILITIES (KR-20)

F-MLCT	.98
T-MLCT	.98
F-IRIT	.98
T-IRIT	.98
SAT Reading	.94
Comprehension	

Summary of Findings

Findings for the data analyses of the hypotheses presented in this chapter as well as additional findings are summarized below:

1. The matching lexical cloze tests showed a significant correlation with the group informal reading inventories, standardized tests and teacher rankings. The correlations between the matching lexical cloze tests and group informal reading inventories ranged from  $r(33) = .44$  to  $.50$ ,  $p \leq .05$ . The correlations between the matching lexical cloze tests and the Stanford Achievement Test, Advanced Battery, Form B Reading Comprehension subtest were  $r(33) = .51$  for the matching lexical cloze test developed from the Finn material and  $r(33) = .60$  for the matching lexical cloze test developed from the Scott-Foresman texts with  $p \leq .05$  in both cases. Correlations with teacher rankings ranged from  $r(35) = .46$  to  $.59$ ,  $p \leq .05$ .

2. Independent, instructional and frustration score ranges for the matching lexical cloze procedures were determined to be 77 to 92, 65 to 76, and 64 percent and below respectively.

3. The matching lexical cloze tests were determined to be highly successful in determining the instructional and frustration levels of students based upon teacher ratings, 70 and 55 percent correct for the instructional level on the sixth and eighth grade



## CHAPTER V

### SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The major purposes of this study were: a) to determine the relationship between the matching lexical cloze procedures, two group informal reading inventory procedures, and a standardized reading test (i.e., The Stanford Achievement Test, Advanced Battery Form B Reading Comprehension Subtest), b) to determine if the presented matching lexical cloze procedure will produce reliable and valid results, and c) to determine when the matching lexical cloze technique can be used by the classroom teacher. In addition, attempts were made: a) to determine comparable independent, instructional, and frustration score ranges for the matching lexical cloze technique with relation to traditional group informal reading inventories and b) to compare the use of a matching lexical cloze test score and teacher ratings of student ability by placing them in independent, instructional, and frustration levels for specific reading passages.

Some of the major terms used in this study are reviewed below:

1. The term matching lexical cloze referred

to tests which systematically deleted only nouns, verbs, and verbals with syntactically correct but semantically implausible distractors listed with the correct responses at the bottom of the page. These tests purportedly measured literal comprehension of specific passages.

2. The term group informal reading inventory referred to tests which were constructed to purportedly measure literal comprehension. Students were required to read passages and answer questions concerning the content of the passages without referring to the passage for information.

3. The term standardized test referred specifically to the Stanford Achievement Test, Advanced Battery Form B Reading Comprehension Subtest, a group administered, nationally normed instrument. The specific subtest used in the research is the section in which students are required to read brief passages and respond to multiple choice questions concerning those passages.

The basic questions were as follows:

1. What were the ranges of scores on a matching lexical cloze procedure for independent, instructional, and frustration levels in light of what is normally considered the range for these levels on a group informal reading inventory?

2. Was a matching lexical cloze procedure a valid and reliable instrument for determining approp-

riate reading levels according to teacher ratings?

3. How did the ranked scores on a matching lexical cloze procedure compare to teacher rankings of students?

4. Did the matching lexical cloze procedure correlate with traditional standardized test techniques?

5. How did test scores from the matching lexical cloze procedures compare with more traditional measures of literal comprehension utilizing the same passage materials as the matching lexical cloze tests, namely group informal reading inventories?

In order to answer these questions eighth grade students from two intact language arts classes in the Lincoln County R-IV Middle School were used. This school district is located in a semi-rural area north of St. Charles County, Missouri. The sample was composed of thirty-five students, eighteen males and seventeen females. These two groups were tested by the same instructor, in the same room, and on the same days.

On the first day, each class was given a sample exercise with the matching lexical cloze procedure and the matching lexical cloze test developed from the Scott-Foresman material. The following day, the matching lexical cloze test developed from the Finn material was administered. The following week, the Stanford Achievement Test Battery Form B was administered as part of the district's testing policies. Fourteen

days after the initial testing the group informal reading inventory based on the Finn material was administered, and seven days later the alternate form group informal reading inventory developed from the Scott-Foresman material was administered.

Hypotheses related to validity comparisons of the matching lexical cloze procedures, group informal reading inventories, and standardized tests were analyzed by means of the Pearson product-moment correlation. The ranking results were analyzed by means of Spearman's Rho ranked correlation formula. Purposes related to development of a range of scores for independent, instructional, and frustration levels using scores on the matching lexical cloze tests and the group informal reading inventories were developed by means of a regression equation and visual inspection.

### Conclusions

The following conclusions are based on the results of testing the hypotheses stated in Chapter 4:

1. The matching lexical cloze tests showed significant positive correlations with group informal reading inventories, standardized tests, and teacher rankings ( $r=.44-.60$ ).
2. Independent, instructional, and frustration score ranges for the matching lexical cloze tests were determined to be 77-92, 65-76, and 64 percent and below respectively.

3. The matching lexical cloze tests seemed highly successful in determining the instructional and frustration levels of students based upon teacher ratings, 55-70 percent correct predictions.

4. All test measures had high KR-20 reliability estimates, .94-.98.

#### Discussion of the Findings

Based on the findings which compared the matching lexical cloze tests and the group informal reading inventories, it appears that both matching lexical cloze tests and both group informal reading inventories were comparable in terms of difficulty. However, the correlations between these measures were much lower than expected,  $r=.44-.50$ . There are several possible reasons for these results: a) student unfamiliarity with the matching lexical cloze format, b) student unfamiliarity with procedures which do not permit reference to material in order to locate responses, and c) the Finn questions, particularly on the sixth grade level passage of the group informal reading inventory.

Only two of the students tested indicated that they had ever had contact with a cloze procedure and that that exposure had been to a traditional every fifth word deletion cloze procedure. None of the students had ever experienced a test format like the matching

lexical cloze instruments used in this research.

During the course of their language arts class, both classes of students had always been permitted to refer to stories or material read in order to obtain correct responses. During the group informal reading inventories, the students were not permitted to refer to the text for answers for the first time in their language arts work. The Finn group informal was the first group informal reading inventory administered to the students. It was noted by this researcher that the students quickly read the first passage but discovered that they had not read carefully enough to respond to the type of questions asked. Reading time on the remaining group informals increased greatly. Students commented that they preferred exercises in which they were allowed to refer to the passages in order to locate responses.

The scores on the group informal reading inventory developed from the Finn material may have been adversely affected by the questions developed and asked on the sixth grade passage (see Appendix C). Several occupations of workers within a mine were contained in the passage (e.g., lamp-man, banksman, engine-man, and onsetter). One of the questions asked was, "What is a man called?" with the only correct response being "banksman". In addition, several specific duties for each occupation were contained in the passage and these

were also asked as information to be answered in the questions developed by Finn for this passage. In the fourth paragraph of the text (see Appendix C), the tunnels and where they lead in the mine are discussed. The text reads, "Some of the tunnels lead to the coal-face. Others lead to store-rooms, workshops, loco-garages,..." One question asked, "Where do some of the tunnels lead?" with the only correct response being, "the coal-face". Several students incorrectly responded "store-rooms or first-aid stations" due to confusion over the interpretation of the indefinite pronoun "some" and the exact, correct response. The information required for correct responses often necessitated recall of detailed material that students were normally accustomed to looking up in the material read from the text.

It appeared that the group informal reading inventories had a tendency to measure immediate recall of specific material contained in the text; whereas, the matching lexical cloze tests appeared to be more related to general reading ability (literal comprehension of material) than recall ability. The consistently low to moderate correlations between the group informal reading inventories and the standardized test seem to reinforce this conclusion.

Even when the matching lexical cloze tests were rescored so that semantically correct responses were accepted rather than exact word substitutions, the

average change in total test scores was only six percentage points (2.5 raw score points) on the tests developed from the Scott-Foresman material and ten percentage points (4 raw score points) on the tests developed from the Finn material. On the second section of the matching lexical cloze test developed from the Finn material, two deletions which may have accounted for the difference in rescored test marks were the words "called" and "named". It is important to note that the rescoring of the matching lexical cloze procedures for semantically correct responses though not exact word substitutions did not significantly change the correlations with the group informal reading inventories,  $r(33) = .42-.50$ . This finding also supports the conclusion that the matching lexical cloze tests measured different aspects of the reading process than did the group informal reading inventories.

Based on the findings relating to teacher ranking of students and the matching lexical cloze procedures, once again the correlations were only moderate,  $r = .46-.59$ ; however, it should be noted that the higher correlations were found on the eighth grade matching lexical cloze passages.

As reported in Chapter 4, score ranges for the independent (77-92 percent), instructional (65-76 percent), and frustration (64 percent and below) levels were determined for the matching lexical cloze

procedure utilizing regression equations and visual inspection. Another factor influencing the determination of the ranges was the nature of the questions in the sixth grade passage of the informal reading inventory developed from the Finn material (see Appendix C). Due to this final factor, the score ranges were based primarily on the text passages developed from the Scott-Foresman material and the regression figures since no student scores were available for the independent level from the matching lexical cloze test based on the sixth grade Finn material or for the total test. Nonetheless, scores on the tests developed from the Finn material were utilized for determination of other level score ranges.

When compared with end-of-year teacher ratings of student reading ability, the matching lexical cloze tests were highly successful in determining the instructional and frustration levels for students. They correctly predicted the instructional and frustration levels 70 percent of the time on the sixth grade passages and 55 and 65 percent of the time on the eighth grade passages. The group informal reading inventories, which took longer to administer, correctly estimated the instructional level 42 and 53 percent of the time and the frustration level 100 and 87 percent of the time on the sixth and eighth grade passages respectively. These facts seem to support the conclusion that the matching

lexical cloze technique can be used most effectively to determine the suitability of particular materials for the purpose of instruction.

#### Recommendations from the Data

1. The matching lexical cloze procedure can be used to develop reliable, consistent, and objective reading comprehension tests.

2. Teachers and reading specialists should work together to develop matching lexical cloze tests in various content areas in order to assess basic reading comprehension of various text materials. Because the matching lexical cloze test is less anxiety producing, easier to construct, easier to score, and more objective than teacher-made instruments, it has several advantages over the traditional cloze procedure or teacher constructed instruments.

#### Recommendations for Further Research

Based upon the limitations and conclusions of this study, the following recommendations for further research are suggested:

1. This study should be replicated with a different population to determine if the results would be comparable.

2. This study should be replicated with a different content area material to determine if results are comparable.

3. A study in which several independent item-writers would produce matching lexical cloze tests using the same reading passages should be conducted to determine the replicability of this procedure.

4. A study also utilizing a standardized group informal reading inventory should be conducted to determine the comparability of the group informal reading inventory scores based on Finn's Algorithm and Bormuth's wh-format.

APPENDIX 1  
 Procedure for preparing  
 the  
 English lexical lists

I. Source Selection Criteria

1. Length

1. unless passages of approximately 200 words, such that there are a total of 40 passages for the list prepared;
2. unless passages are only of about 100 words of length, unless it is possible to have passages of 200 words and passages of 100 words of length, unless it is possible to have passages of 200 words and passages of 100 words of length, unless it is possible to have passages of 200 words and passages of 100 words of length.

2. Qualifications

1. passages must be carefully selected following the above 10 suggested guidelines;
2. lists are prepared in interlinear form with word examples, illustrations, etc., with examples of use of the word together in the native language;
3. the following may be included in each list:
  - a. translations (head)
  - b. reference to parts, illustrations, diagrams, etc.
  - c. examples and problems
  - d. context material or headings

APPENDICES

II. Listing the words

1. Rules for listing

1. only 40 words only words, verbs, and nouns, (e.g., particles, gerunds, and infinitives).
2. Do not list the following:
  - a. function words (conjunctions, prepositions, interjections, auxiliary verbs)
  - b. adjectives
  - c. adverbs
  - d. pronouns
  - e. proper nouns or names (capitalized in the first letter (e.g., Mary))
  - f. numbers that provide no extra help (e.g., three, twenty, forty).

## APPENDIX A

### Procedure for Developing the Matching Lexical Cloze Tests

- I. Passage Selection Criteria
  - A. Length
    1. Select passages of approximately 200 words, such that there are a total of 40 deletions for the test passages.
    2. Select passages that have at least 100 words of continuous prose preceding the test passage.
    3. Begin the test passage with the beginning of a paragraph (two undeleted sentences) and end the passage at the end of the sentence containing the final deletion.
  - B. Quality-coherence
    1. Passages must be coherent-one sentence following another in connected discourse.
    2. When the discourse is interspersed with many examples, problems, illustrations, etc., such examples may be drawn together to form unified passages.
    3. The following may be deleted to meet the criterion of coherence:
      - a. transitional phrases
      - b. references to charts, illustrations, diagrams, etc.
      - c. examples and problems
      - d. chapter subtitles or headings
- II. Clozing the Passage
  - A. Rules for Deletions
    1. Cloze or delete only nouns, verbs, and verbals, (e.g., participles, gerunds, and infinitives).
    2. Do not cloze or delete:
      - a. function words (conjunctions, prepositions, interjections, auxiliary verbs)
      - b. adjectives
      - c. adverbs
      - d. pronouns
      - e. proper nouns or nouns capitalized in a phrase (e.g., Group A)
      - f. numbers that function as nouns (e.g., first, second, forty-five)

- g. hyphenated words or two words that work together but are not hyphenated (e.g., decision making)
  - h. Arabic or Roman numerals
  - i. abbreviations
  - j. phonemes such as aw, oo, ah
  - k. foreign words (see Webster's Seventh New Collegiate Dictionary)
  - l. any form of the verb to be (e.g., is, are, was, were, be, am)
  - m. verbs-became, seems, means (difficult to find grammatically plausible distractors)
  - n. idioms (i.e., idiosyncrasies of our language, and, above all those phrases which are verbal anomalies, which transgress either the laws of grammar or the laws of logic; a small group of words expressing a single notion or entering with some degree of unity into the structure of the sentence
    1. adverbial phrases consisting of terse adverbial phrases formed by a preposition with a noun or adjective (e.g., at hand, by chance, for example, in fact, of course, in general, for good, on hand, etc.)
    2. verb particles--verbs followed by prepositions or by prepositions used as adverbs (e.g., added up, keep down, set up, etc.)
    3. phrases where two words are used together as emphasis (e.g., hammer and tongs, heart and soul, etc.)
    4. meaning of the phrase is other than the meaning of the words that compose it (e.g., an axe to grind, beat around the bush, etc.)
    5. other phrases (a great deal, for the sake of, to keep in mind, in dispute of, to the extent that)
  - o. nouns used in phrases that function as conjunctions (e.g., due to the fact that, on the contrary, etc.)
- B. Procedure for Deletions
1. Delete every fifth word as often as possible, i.e., leave four words between deletions)
  2. If the fifth word can be deleted, circle it.
  3. If not, check to see if the sixth word can be deleted, if so, circle it.
  4. If not, check to see if the fourth word can be deleted, if so, circle it.

5. If not, continue forward at the seventh word until a word can be deleted or clozed.
6. It is permissible to leave as many as 11 words between deletions but in no more than four instances per passage, can there be more than eight words between deletions; if this condition is not met, select another passage to cloze.
7. Continue to use this system of deletions until 40 words are deleted or clozed.

### III. Developing Distractor Lists

#### A. Source

1. Begin with the sentence immediately preceding the initial sentence of the test passage.
2. Beginning with the final word of that sentence, follow the above mentioned deletion or cloze procedure in reverse. (e.g., The horse ran swiftly to the barn. The fifth word is "ran" and qualifies for deletion.; The horse swiftly ran to the barn. The sixth word is "horse" and qualifies for deletion.)
3. Underline the words selected as distractors.
  - a. Count as one word--proper nouns or adjectives, dates, other symbols, numbers, hyphenated words (e.g., Joe, 1945, IRA, & post-test).
  - b. Do not select nouns or verbs of the types included in the previous discussion as the types of words that would not be deleted or clozed in the original text.
  - c. Do not select as distractors words that may be synonymous with a deleted word (e.g., quick=fast, drunk=inebriated).
  - d. Do not select semantically plausible distractors within the context of the entire passage intact.

Examples:

"Six more weekends (years) would pass by before I see him again."

"He liked football (hockey) because it was a very physical sport."

- e. Do not select grammatically implausible distractors.
- f. Select distractors so that they agree in number and tense of the deleted words of the passage. If this is not possible due to grammatical structures of the paragraphs, suitable form distractors



## APPENDIX B

## Matching Lexical Cloze Test Forms

SAMPLE EXERCISE

**DIRECTIONS:** Read the following passage. You will notice that words have been left out and replaced with numbered blanks. Below the story is a list of words that might fit into the blanks. Some might not be used at all. Only one word fits correctly in each blank space. On a separate sheet of paper number from one to five (1-5). Write your answers on this separate sheet. Make sure your numbers and answers match the numbered blanks from the passage.

Once there was a school and it was closed when they got there. It was closed because the crop was ripe. A crop vacation, 1 called this, and everybody 2, young ones and grown-ups and old 3. Everybody except, of course, Princess Anne. Over in Louisiana she 4 by herself in a fruit 5 at the end of the strawberry rows and sucked her thumb.

sat

folks

picked

girls

families

school

people

crate

## T-MIC 6

READ THE FOLLOWING PASSAGE. The words at the bottom of the page are the possible choices to fill in the blanks. Some of the words will not be used at all. Write the correct answers on the answer sheets provided. These words are for numbers 1-20 only. Be sure the number on the answer sheet matches the blank.

Few people have seen a tiger stalk, attack, and kill an animal. Tigers are shy. Whenever possible, they 1 through dense thickets. To 2 tigers, I have traveled both 3 and night through the 4 and meadows. Sometimes 5 at places where I 6 them to pass. The best 7 to watch a 8 is at its 9. If the kill is large, the 10 may remain beside it for several 11, until the last scrap of 12 is eaten.

I watched the 13 from blinds, or hiding 14. My blinds would be a small wooden 15, a thin screen of 16 and leaves, and sometimes my 17. One mother tiger with 18 grew used to my being near her 19. They sometimes just 20 about one hundred feet while I walked back and forth.

day	study	way	days	meat
expected	animals	places	car	forests
patterns	kill	wildlife	tiger	move
cubs	tell	retreated	kills	tigers
shelter	tiger	stripes	waited	grass
		recognize		

READ THE FOLLOWING PASSAGE. The words at the bottom of the page are the possible choices to fill in the blanks. Some of the words will not be used at all. Write the correct answers on the answer sheets provided. These words are for numbers 21-40 only. Be sure the number on the answer sheet matches the blank.

Lope saw that if he remained within reach of his wife's cousin, the tax collector, he would soon be a poor man once again. So he decided to pack up his remaining treasure and move to another part of Spain. For this purpose, he 21 a stout mule and 22 it in the gloomy 23 beneath the Tower of the Seven Stories. This was the very 24 from which it was 25 that Belludo, the headless 26, raced forth every 27 with a pack of 28 at his heels to 29 the streets of Granada. Lope Sanchez did not 30 this story. But he 31 use of it, 32 that no one would be likely to 33 into the vaults beneath the Tower of the Seven Stories. He 34 his family off during the 35 with orders to 36 for him in a distant 37. Then, as night came, he 38 his treasure to the 39. When he had 40 up his mule, he led it forth through the dark avenues.

loaded	wait	place
said	carried	roam
village	tower	sent
bought	remained	day
believe	night	made
tethered	pry	learned
hounds	vault	reminded
prison	apartments	knowing
	paid	horse

relay, then near the top of the opening shaft, the  
 the used air and the fire-damp through ledge which the  
 cause fresh air to be pulled in at the bottom and  
 passed through the mine. He noticed that \_\_\_\_\_ had a  
 mine, and great \_\_\_\_\_ is taken out to \_\_\_\_\_ quarries.  
 a lamp which had a \_\_\_\_\_, but which did not set fire  
 to the \_\_\_\_\_ was suggested to be \_\_\_\_\_ of the  
 called the very Safety lamp. The \_\_\_\_\_ to be a lamp  
 \_\_\_\_\_ to be taken as much as \_\_\_\_\_ of the mine. The  
 lamp is \_\_\_\_\_ to a battery which is \_\_\_\_\_ to the shaft's  
 belt. By \_\_\_\_\_ the lamp that air \_\_\_\_\_ the mine, the  
 the lamp was \_\_\_\_\_ the only mine air \_\_\_\_\_  
 been today, only the water \_\_\_\_\_ to \_\_\_\_\_ the  
 electric light.

The \_\_\_\_\_ are taken to and the \_\_\_\_\_ is  
 \_\_\_\_\_ there are like \_\_\_\_\_ and will not set fire to the lamp.

## F-MLC 6

READ THE FOLLOWING PASSAGE. The words at the bottom of the page are the possible choices to fill in the blanks. Some of the words will not be used at all. Write the correct answers on the answer sheet provided. These words are for numbers 1-20 only. Be sure to match the number on the answer sheet with the number of the blank you are answering.

Today, fans near the top of the upcase shaft suck out the used air and the firedamp through large funnels. This causes fresh air to be pulled in at the downcast and passed through the mine. No matches are 1 down a mine, and great 2 is taken not to 3 sparks. A lamp which had a 4, but which did not set fire to the 5, was a great help to 6. It was called the Davy Safety Lamp. The 7 also has a lamp 8 to his helmet so that he can 9 to work. The lamp is 10 by a battery which is 11 on the miner's belt. By 12 the lamps that are 13 out each day, the lamp-man 14 how many miners are down the 15. Even today, only the main 16 in mines are 17 by electric lights.

The 18 are taken to and from the pit-bottom in 19. These are like 20 and each cage has two decks.

make	fixed	mines
get	miner	counting
mine	knows	taken
miners	work	flame
lifts	explode	given
care	lit	tunnels
see	cages	called
candles	gas	fixed
	lit	miners

## F-MLC 8

READ THE FOLLOWING PASSAGE. The words at the bottom of the page are the possible choices to fill in the blanks. Some of the words will not be used at all. Write the correct answers on the answer sheets provided. These words are for numbers 21-40 only. Be sure the number on the answer sheet matches the blank.

One of the most famous of all pirates, whose name is known to every schoolboy, was Captain Kidd. Actually, he could scarcely be called a pirate at all. The son of a 21, he was born in Scotland in 1655 and was 22 to command what was 23 a privateer when he was in his early thirties. A 24 was ship given legal 25 to search out and 26 the ships of the 27 or queen's enemies. It was 28 for by perfectly respectable 29 who took a 30 of the profits of the 31 either by the 32 of captured ships of the 33 taken out of them.

In 1695, William III, 34 of England, gave Captain Kidd, whom he 35 his beloved friend William Kidd, the 36 of a privateer 37 Adventure with orders to 38 piracy off the 39 of America. Captain Kidd was unlucky. He 40 some ships in the service of France, which he had a perfect right to do as Britain and France were at war at the time.

voyage	coast	sale	privateer
chest	share	parson	pieces
authority	king	remained	commissioned
persons	captured	kings'	command
search	suppress	called	called
treasure	occupation	destroy	
spot	named	paid	

Today, Jane was the top of the mine shaft. The wind blew and the fire burned through the pipes. The mine was full of smoke and the air was hot. The men were working hard and the noise was loud. The mine was a dangerous place and the men were in great danger. The mine was a source of wealth and power and the men were fighting for it. The mine was a symbol of the industrial revolution and the men were the workers who made it possible.

The mine was a complex system of tunnels and shafts. The men worked in the dark and the air was thick with dust. The men were paid very little and they worked long hours. The mine was a place of suffering and death. The men were treated like machines and they were used to produce wealth for a few men. The mine was a symbol of the exploitation of the working class.

All the men were poor and they were all struggling to survive. The mine was a place of hope and despair. The men were working for a better life but they were getting nowhere. The mine was a place of struggle and the men were fighting for their rights. The mine was a symbol of the fight for social justice.

The men were tired and they were all looking for a way out. The mine was a place of no return and the men were trapped. The mine was a place of death and the men were dying. The mine was a symbol of the end of the world.

## APPENDIX C

### Informal Reading Inventory Test Forms

DIRECTIONS...F-IRI 6 Text

Coal

There is great danger in mines from gas. The gas is called firedamp and it explodes if a flame gets near it. The early miners had to use candles to be able to work. Sometimes the candles set the gas alight and caused explosions.

Today, fans near the top of the upcase shaft suck out the used air and the firedamp through large funnels. This causes fresh air to be pulled in at the downcase and passed through the mine. No matches are taken down a mine and great care is taken not to make sparks. A lamp which had a flame, but did not set fire to the gas, was a great help to miners. It was called the Davy Safety Lamp. The miner also has a lamp fixed to his helmet so that he can see to work. The lamp is lit by a battery which is fixed on the miners belt. By counting the lamps that are given out each day, the lamp-man knows how many miners are down the mine. Even today, only the main tunnels in mines are lit by electric lights.

The miners are taken to and from the pit-bottom in cages. These are like lifts and each cage has two decks. There are two cages in each shaft and when one cage is at the top, the other is at the bottom. When the top cage is loaded, a man called the banksman signals to the engine-man at the winding house. The engine-man then waits for a signal from the onsetter at the bottom cage. After he has both signals, the engine-man starts the winding gear. The cages never carry men and coal together. The banksman and onsetter signal to the winding engine-man what the cages are carrying. When the cages are carrying men, the winding gear is run slower than if coal is being lifted.

At the pit-bottom, many tunnels lead off into the mine. The roofs of the tunnels are held up by strong steel girders. Some of the tunnels lead to the coal-face. Others lead to store-rooms, workshops, loco-garages, the first-aid room and the fire-station. Narrow railway lines lead to the coal-face. This may be as far as six miles away. If the coal-face is far away, the miners ride on railcars pulled by diesel locomotives.

Near the coal-face, the roof is lower and the miners may have to bend down as they walk. The only lights are those on their helmets. The coal-face is a shining layer

of coal trapped between layers of rock. Near the coal-face are the machines the miners use to cut out the coal. One machine is called a coal-cutter and works by an electric motor. The coal-cutter has a long cutting arm with steel teeth. The cutting arm is pushed into a slot about five feet deep along the coal-face. After the coal-cutter has moved along, other miners drill deep holes into the coal-face. The holes are for explosive charges which will break down the coal-face.

RAISE YOUR HAND AND YOU WILL BE GIVEN A SHEET OF QUESTIONS WITH AN ANSWER SHEET ATTACHED. PUT YOUR NAME ON THE ANSWER SHEET. YOUR ANSWERS WILL BE ONE OR TWO WORD ANSWERS. WHEN YOU HAVE FINISHED WITH THE ANSWERS TO THOSE QUESTIONS, RAISE YOUR HAND.

1. What is a coal-cutter?
2. What is a coal-cutter used for?
3. RAISE YOUR HAND AND YOU WILL BE GIVEN A SHEET OF QUESTIONS WITH AN ANSWER SHEET ATTACHED. PUT YOUR NAME ON THE ANSWER SHEET. YOUR ANSWERS WILL BE ONE OR TWO WORD ANSWERS. WHEN YOU HAVE FINISHED WITH THE ANSWERS TO THOSE QUESTIONS, RAISE YOUR HAND.
4. As it moves along the coal-cutter cuts a slot.
5. What way do we get an idea of the size of a slot?
6. Where is the coal-cutter used?
7. Where is some of the coal-cutter used?
8. What will explosive charges break down?
9. To what is the explosive and smaller signal used? The coal-cutter is carrying?
10. What is a striking layer of coal-trapped between layers of rock?
11. When the top layer is loaded and a man signals him, who else waits for a signal from the assistant?
12. After he has both signals and starts the striking gear?
13. If a flame gets near it what happens?
14. Where is the coal-cutter?
15. When the coal-cutter is loaded what does a man signal?
16. After the coal-cutter has moved along what do other miners drill into the coal-face?
17. What are the holes used for?
18. Where are the explosive charges?
19. What is a coal-cutter?

## F-IRI 6 Questions

## COAL

1. What are like lifts?
2. What has a long cutting arm?
3. What is the gas called?
4. What never carries men and coal together?
5. What are miners taken to and from the pit-bottom in?
6. Today what do fans suck out through large funnels?
7. The miners ride on rail cars if what is far away?
8. The winding gear is run slower than if coal is being lifted when what are carrying men?
9. What is one machine called?
10. As it moves where does the cutter cut a slot?
11. What may be as far as six miles away?
12. Where do railway-lines lead?
13. Where do some of the tunnels lead?
14. What will explosive charges break down?
15. To whom do the banksman and onsetter signal what the cages are carrying?
16. What is a shining layer of coal trapped between layers of rock?
17. When the top cage is loaded and a man signals him, who then waits for a signal from the onsetter?
18. After he has both signals who starts the winding gear?
19. If a flame gets near it what explodes?
20. Where is the roof lower?
21. When the top cage is loaded whom does a man signal?
22. After the coal-cutter has moved along where do other miners drill deep holes?
23. What are in each shaft?
24. Where are the machines?
25. What is a man called?

## PIRATES

When one sets sail today on a cruise in the Mediterranean or down the coast of Africa or across to the West Indies, it is not necessary to keep a sharp look-out for pirates. Thanks largely to the British navy, pirates are no longer amongst the dangers to be met on the high seas. But if you had embarked on any of these voyages in the time of Elizabeth I, or for more than two hundred years afterwards, somebody on the ship would have been keeping a very sharp look-out indeed.

Between the years 1550 and 1750, there were literally hundreds of pirates roving the seas, boarding and sinking ships, cutting throats and then retiring to some snug anchorage, often in the West Indies, to enjoy or to bury the treasure they had taken. Although it was the British navy which played the largest part in stopping piracy, most of the pirates were from the British Isles. It is true that many of them were caught and hanged, others buried their treasure on some remote island hoping to come back for it later. Some of these buried hoards are still there and even today expeditions are organized to search for them. It is a romantic occupation but one which is rarely successful. The lure of the ancient sea-stained map remains with the cross marking the spot where the old iron-bound chest full of pieces of eight is waiting.

One of the most famous of all pirates, whose name is known to every schoolboy, was Captain Kidd. Actually, he could scarcely be called a pirate at all. The son of a parson, he was born in Scotland in 1655 and was commissioned to command what was called a privateer when he was in his early thirties. A privateer was a ship given legal authority to search out and destroy the ships of the kings' or queens' enemies. It was paid for by perfectly respectable persons who took a share of the profits of the voyage either by the sale of captured ships or the treasure taken out of them.

In 1695, William III, king of England, gave Captain Kidd, whom he called his beloved friend William Kidd, the command of a privateer named Adventure with orders to suppress piracy off the coast of America. Captain Kidd was unlucky. He captured some ships in the service of France which he had a perfect right to do as Britain and France were at war at the time. But when he arrived at New York with ninety bars of gold, he was arrested as a pirate and executed not for piracy but for killing a mutinous member of his crew by hitting him over the head with a bucket. In the charge sheet, the bucket is

carefully described as being of the value of eight pence.  
 RAISE YOUR HAND WHEN YOU HAVE FINISHED READING.

1. The work is carefully reviewed as being of the value of eight pence.
2. What did pirates carry on some island before they left?
3. Describe the ship.
4. What is still on some remote island?
5. When was the first call on a whaler in the Arctic region or down the coast of Alaska or to the West Indies in the last century to help with the pirates?
6. When did William III give command to a privateer in 1692?
7. What was the name of the privateer?
8. When did William III call his beloved friend?
9. If you had returned on any of these voyages in the time of Queen Elizabeth I or the first half of the last century what would you have seen?
10. When Captain Cook arrived at New York what was the name of the ship?
11. What pirates were from the British Isles although it was the British navy that played them?
12. Who was one of the most famous pirates of all times?
13. How was he caught?
14. What was his name?
15. Between the years of 1550 and 1750 after roving the seas boarding and sinking ships and getting various pirates retired where to enjoy or bury treasure?
16. What was a ship given legal authority to search for and destroy ships of Kings and Queens elsewhere?
17. What was Captain Cook's name?
18. What was the name of Captain Cook's ship?
19. When did William III give Captain Cook command in 1769?

## F-IRI 8 Questions

## PIRATES

1. The bucket is carefully described as being of what value in the charge sheet?
2. What did pirates bury on some island hoping to come back for it later?
3. Describe the chest.
4. What is still on some remote island?
5. When one sets sail on a cruise in the Mediterranean or down the coast of Africa or to the West Indies it is not necessary to keep what for pirates?
6. Whom did William III give command of a privateer in 1695?
7. What was named Adventure?
8. Whom did William III call his beloved friend?
9. If you had embarked on any of these voyages in the time of Queen Elizabeth I or for more than two hundred years what would someone have been keeping?
10. When Captain Kidd arrived at New York with ninety bars of gold what was he executed for?
11. Most pirates were from the British Isles although it was the British navy that played what?
12. Who was one of the most famous pirates of all times?
13. Who was unlucky?
14. What had pirates taken?
15. Describe the map.
16. Between the years of 1550 and 1750 after roving the seas boarding and sinking ships and cutting throats pirates retired where to enjoy or bury treasure?
17. What was a ship given legal authority to search out and destroy ships of kings and queens enemies?
18. When was Captain Kidd born in Scotland?
19. What was a member of Captain Kidd's crew?
20. What did William III give Captain Kidd orders to suppress in 1695?

## T-IRI 6 Text

In studying wild animals, it is helpful to be able to recognize individuals. The black stripes on each tiger's face form different patterns. So in time I could tell one tiger from another by its markings.

Few people have seen a tiger stalk, attack, and kill an animal. Tigers are shy. Whenever possible, they move through dense thickets. To study tigers, I traveled both day and night through forests and meadows. Sometimes I waited at places where I expected them to pass. The best way to watch a tiger is at its kill. If the kill is large, the tiger may remain beside it for several days, until the last scrap of meat is eaten.

I watched the tigers from blinds, or hiding places. My blinds would be a small wooden shelter, a thin screen of grass and leaves, and sometimes my car. One mother tiger and her cubs grew used to my being near her kills. They sometimes just retreated about one hundred feet while I walked back and forth to the blind. Most tigers in this park had not been shot at or harmed by people for many years. Several of them quickly became used to me.

A tiger eats whatever animals it can catch. This includes birds, lizards, frogs, snakes, and locusts. It even eats grass and earth. We don't know why.

Today, much of the wildlife in India that tigers usually hunted for food has been killed off by people. So tigers in most forests kill tame animals--usually cattle and buffalo.

Tigers hunt mostly at night. They walk through their range, or hunting area, searching for food. At Kanha I saw them often on forest roads. There they could travel easier, quieter, and faster than in the high grass or dead leaves of the forest. The bottoms of the many dry stream beds were also much used by tigers for traveling. Tigers may travel ten to twenty miles a night in their search for food. Sometimes they travel thirty miles.

The tiger spots its prey usually by sight and -- at night -- also by hearing. I saw no signs that tigers use their noses to find live animals.

## T-IRI 6 Questions

1. What forms different patterns on each tiger's face?
2. Who has seen a tiger stalk and kill an animal?
3. Where do tigers move whenever possible?
4. What did the writer of this story travel through to study tigers?
5. What did the writer do at places he expected tigers to pass?
6. What is the best way to watch a tiger?
7. Where might a tiger stay for several days?
8. Why would the tiger stay there?
9. What are the hiding places for watching tigers called?
10. What had not happened to the tigers in the park for years?
11. What kinds of animals do tigers eat? (not a specific kind)
12. What has killed much of the wildlife in India?
13. What do tigers in India kill today?
14. When do tigers hunt?
15. What is the hunting area of tigers called?
16. Why were tigers often seen on the roads?
17. What else was used by tigers for traveling purposes?
18. How far may tigers travel in their search for food?
19. How does a tiger spot its prey?
20. The writer said he didn't see that tigers use what to find live animals?

## T-IRI 8 Text

The tax collector pondered for a long time. "It may be that if your husband pays a tax on this treasure he will be allowed to keep it. For the time being you have to give me only two bags of gold to pay the tax for Finding Buried Treasure. I will look up the law on the matter and arrange everything."

When Lope's wife heard this, she was pleased to think she could save her husband from prison so cheaply. She brought forth two bags of gold at once and gave them to her cousin to pay the tax for Finding Buried Treasure.

But when Lope Sanchez heard of this, he almost went mad. "What will become of us?" he cried.

His wife reminded him of the countless wealth that still remained and of how she had saved him from prison. It seemed, however, that there were many, many taxes to be paid on this treasure. Each day the cousin returned to Lope's apartment with a fresh tax he had learned of. There was a tax for Working after Sundown; there was a tax on Enchanted Jewels; there was a tax on Spellbound Gold; there was even a tax on Exploring Underground Vaults!

Lope saw that if he remained within reach of his wife's cousin, the tax collector, he would soon be a poor man once again. So he decided to pack up his remaining treasure and move to another part of Spain. For this purpose, he bought a stout mule and tethered it in the gloomy vault beneath the Tower of the Seven Stories. This was the very place from which it was said that Belludo, the headless horse, raced forth every night with a pack of hounds at his heels to roam the streets of Granada. Lope Sanchez did not believe this story. But he made use of it, knowing that no one would be likely to pry into the vaults beneath the Tower of the Seven Stories.

## SELECTED BIBLIOGRAPHY

- Aaron, Ira E.; Gray, Gordon W.; Johns, Annette L.; Johns, Jerry L.; Koke, Harold J.; Koke, Bena A.; Manning, John C.; Rainwater, Jerry A.; Smith, Richard G.; and Tierney, Robert J.; Batter Up! Scott, Foresman Basics in Reading Series. Glenview, Illinois: Scott, Foresman and Co., 1978.
- Aborn, M.; Rubenstein, H.; and Sterling, T.D. "Sources of Contextual Constraint upon Words in Sentences." Journal of Experimental Psychology. 3 (1959): 117.
- Baldauf, R. B., Jr.; Dawson, R. L. T.; Prior, John; and Propst, I. K. Jr.; "Can Matching Cloze Be Used with Secondary ESL Pupils?" Journal of Reading. 23 (1980): 435-440.
- Baldauf, R. B. Jr.; Propst, I. K. Jr.; "Matching and Multiple-Choice Cloze Tests." Journal of Educational Research, Vol. 72, 6 (July-August, 1979): 321-326.
- Betts, E. A., Foundations of Reading Instruction. New York: American Book Co., 1946.
- Bormuth, J. R. "Cloze Tests as Measures of Readability and Comprehension Ability." Ph.D. dissertation, University of Indiana, 1962.
- \_\_\_\_\_. "Experimental Applications of Cloze Tests." Paper presented at the International Reading Association Conference, 1964.
- \_\_\_\_\_. "Mean Word Depth as a Predictor of Comprehension." California Journal of Educational Research, 15 (1964).
- \_\_\_\_\_. "Optimum Sample Size and Cloze Test Length in Readability Measurement." Journal of Educational Measurement, 2 (1965).
- \_\_\_\_\_. "Validities of Grammatical and Semantic Classifications of Cloze Scores." Paper presented at the International Reading Association Conference, 1965.

- \_\_\_\_\_. "Readability: A New Approach." Reading Research Quarterly, 1 (1966).
- \_\_\_\_\_. "Comparable Cloze and Multiple-Choice Comprehension Test Scores." Journal of Reading, 10 (1967).
- \_\_\_\_\_. On the Theory of Achievement Test Items. Chicago: University of Chicago Press, 1970.
- \_\_\_\_\_. "Reading Literacy: Its Definition and Measurement." Reading Research Quarterly, 9 (1974): 7-66.
- Coleman, E. B., and Blumenfeld, J. P. "Cloze Scores of Nominalizations and their Grammatical Transformations Using Active Verbs." Psychological Reports, 13 (1963): 651-654.
- Cranney, A. G. "The Construction of Two Types of Cloze Reading Tests for College Students." Journal of Reading Behavior, 51 (1972073): 60-64.
- Dale, E., and Chall, J. "A Formula for Predicting Readability." Educational Research Bulletin, 27 (1948): 11-28.
- Fillenbaum, S.; Jones, L. V.; and Rapoport, A. "The Predictability of Words and their Grammatical Classes as a Function of Rate of Deletion from a Speech Transcript." Journal of Verbal Learning and Verbal Behavior, 2 (1965): 186-194.
- Finn, P. J. "An Item Writing Algorithm for Written English." Ph. D. dissertation, University of Chicago, 1973.
- Friedman, M. "The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at the University of Florida." Ph. D. dissertation, University of Florida, 1964.
- Fry, E. "A Readability Formula that Saves Time." Journal of Reading, 11 (1968): 513-16.
- Greene, F. P. "A Modified Cloze Procedure for Assessing Adult Reading Comprehension." Ph. D. dissertation, University of Michigan, 1964.
- \_\_\_\_\_. "Modification of the Cloze Procedure and Changes in Reading Test Performances." Journal of Educational Measurement, 2 (1965): 213-217.

- Guthrie, J. T.; Seifert, M.; Burnham, N. A.; and Caplan, R. I. "The Maze Technique to Assess, Monitor Reading Comprehension." The Reading Teacher, 28 (1974): 161-168.
- Hafner, L. E. "Cloze Procedure." Journal of Reading, 9 (1966): 415-421.
- Heaney, L. The Cloze Procedure. Fond du Lac, Wisc.: ERIC Document Reproduction, Ed 151 762, 1978.
- Jenkinson, M. E. "Selected Processes and Difficulties in Reading Comprehension." Ph. D. dissertation, University of Chicago, 1957.
- Jones, M. B., and Pikulski, E. C. "Cloze for the Classroom." Journal of Reading, 17 (1974): 432-438.
- Louthan, V. "Some Systematic Grammatical Deletions and their Effects on Reading Comprehension." English Journal, 54 (1965): 295.
- Luke, M. "Form Class and Cloze Procedure." Manuscript, Language Development Program, Center for Human Growth and Development, University of Michigan, 1964.
- McLeod, J., and Anderson, J. "Readability Assessment and Word Redundancy of Printed English." Psychological Reports, 19 (1966): 35-38.
- Neville, M. H., and Pugh, A. K. "Context in Reading and Listening: Variations in Approach to Cloze Tasks." Reading Research Quarterly, 12 (1976-77): 15-31.
- O'Reilly, R. P.; Schuder, S. J.; Kidder, S. J.; Salter, R.; and Mayford, P. D. The Validation and Refinement of Measures of Literal Comprehension in Reading for Use in Policy Research and Classroom Management, Albany, New York: ERIC Document Service Reproduction, Ed 133 363, 1976.
- O'Reilly, R. O., and Streeter, R. C. "Report on the Development and Validation of a System for Measuring Literal Comprehension in a Multiple-Choice Cloze Format: Preliminary Factor Analytic Results." Journal of Reading Behavior, 9 (1977): 45-69.
- Pikulski, J. J.; and Pikulski, E. C. "Cloze, Maze and Teacher Judgment." The Reading Teacher, 30 (1977): 766-770.

- Potter, T. C. A Taxonomy of Cloze Research, Part 1: Readability and Reading Comprehension. Washington, D. C.: ERIC Document Reproduction, ED 035 514, 1970.
- Rankin, E. F. "An Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension." Ph. D. dissertation, University of Michigan, 1957.
- \_\_\_\_\_. "The Cloze Procedure--Its Validity and Utility." In Eighth Yearbook of the National Reading Conference. Causey, O. S., and Eller, W. (Eds.). National Reading Conference, Inc., 8 (1959): 131-144.
- \_\_\_\_\_. "Research Design and the Cloze Procedure." Paper presented at the International Reading Association Convention, 1966.
- \_\_\_\_\_. "The Cloze Procedure Revisited." In Interaction: Research and Practice for College-adult Reading, Twenty-third Yearbook of the National Reading Conference. Nacke, P. L. (ed.). Clemson, South Carolina: National Reading Conference, Inc., 1974.
- Rankin, E. F., and Culhane, J. W. "Comparable Cloze and Multiple-choice Comprehension Test Scores." Journal of Reading, 13 (1969-70): 193-198.
- Rocchio, D. J. "The Validity of Four Non-Traditional Measures of Reading Comprehension with Community College Students." Ed. D. dissertation, University of Missouri-St. Louis, 1979.
- Ruddell, R. B. "A Study of the Cloze Comprehension Technique in Relation to Structurally Controlled Reading Material." In Improvement of Reading Through Classroom Practice. Figurel, J. A. (ed.). International Reading Association Conference. 9 (1964): 298-306.
- \_\_\_\_\_. "The Effect of Oral and Written Patterns of Language Structure on Reading Comprehension." Reading Teacher, 18 (1965): 270.
- Schneyer, J. W. "Use of the Cloze Procedure for Improving Reading Comprehension." Reading Teacher, 19 (1965): 174-179.
- Smith, N., and Zinc, A. "A Cloze-based Investigation of Reading Comprehension as a Composite of Sub-skills." Journal of Reading Behavior, 9.4 (1977): 395-398.

Smith, R. G., and Tierney, R. Racing Stripes, Scott, Foresman Basics in Reading Series. Glenview, Illinois: Scott, Foresman and Co., 1978.

Taylor, W. L. "Cloze Procedure: A New Tool for Measuring Readability." Journalism Quarterly, 30 (1953): 415-433.

\_\_\_\_\_. "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude." Journal of Applied Psychology, 41 (1957): 19-26.