

DETERMINING INSTRUCTIONAL READING LEVEL:
AN INVESTIGATION OF THE RELATIONSHIP AMONG
STANDARD CLOZE TESTS, MULTIPLE CHOICE CLOZE TESTS
AND THE INFORMAL READING INVENTORY

By

Susan Lubet Homan

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL PULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1978

Edw

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to all members of my committee.

Dr. Powell, chairman of the advisory committee has been helpful in several ways. His support and confidence in my abilities have been of tremendous help. His guidance through the various stages of writing the dissertation has been invaluable. More importantly, his example and attitudes as a professional will have a lifelong influence on the quality of my professional life.

My special thanks also to Dr. Linda Crocker for her assistance and advice. It was especially appreciated that she was always willing to help me though often surrounded by other pressing responsibilities.

Both Dr. William Ware and Dr. Lawrence Smith continually offered advice and support. Dr. Ware was of special help with the statistical analysis.

My thanks go to Dr. Lewis who offered his assistance and willingly served on my committee though his commitments were already overwhelming.

Many thanks to Dr. Maria Llabre whose help with the data analysis was lifesaving, and whose special friendship was invaluable.

Also appreciated was the help of Dr. Jeri Benson. Her willingness to help with the finer details of the data analysis, regardless of other commitments, was especially helpful.

To Mary Garcia, I wish to express thanks for typing all of the early drafts, and for her continual encouragement.

I owe a great deal to Richard, my husband, for his compassionate understanding, his willingness to help, and his never faltering faith in my ability.

To my parents, I owe an exceptional debt of thanks for providing an atmosphere that encouraged my continuing education, and for their continuous and unswerving support and confidence in all my life decisions.

TABLE OF CONTENTS

| | PAGES |
|----------------------------------------------------------------------------|-------|
| ACKNOWLEDGEMENTS | 111 |
| ABSTRACT | vii |
| CHAPTER I INTRODUCTION | 1 |
| Statement of the Problem | 5 |
| Objectives and Research Questions | 6 |
| Definition of Terms | 9 |
| Limitations of the Study | 12 |
| CHAPTER II REVIEW OF RELATED LITERATURE | 14 |
| History and Development | 14 |
| Methodological Considerations | 16 |
| Reliability | 21 |
| Criteria | 22 |
| Measurement of Reading Comprehension | 24 |
| Summary of Research | 29 |
| CHAPTER III PROCEDURES AND METHODOLOGY | 31 |
| Sample | 31 |
| Instruments Used | 32 |
| Instrument Development - Original Passage | 33 |
| Other Instruments | 37 |
| Administration | 41 |
| Method of Analysis | 41 |
| CHAPTER IV RESULTS, DISCUSSION AND THEORETICAL CONSIDERATIONS | 47 |
| Results and Discussion | 47 |
| Theoretical Considerations | 85 |
| CHAPTER V SUMMARY AND CONCLUSIONS | 89 |

| | PAGES |
|------------------------------------------------------|-------|
| APPENDIX A DIRECTIONS FOR THE CLOZE TESTINGS | 96 |
| APPENDIX B STANDARD CLOZE PASSAGES | 99 |
| APPENDIX C HUMAN MULTIPLE CHOICE CLOZE PASSAGES . . | 107 |
| APPENDIX D KIDDER MULTIPLE CHOICE CLOZE PASSAGES . . | 120 |
| APPENDIX E INFORMAL READING INVENTORY | 129 |
| LIST OF REFERENCES | 145 |
| BIOGRAPHICAL SKETCH | 150 |

Abstract of Dissertation
Presented to the Graduate Council
of the University of Florida
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

DETERMINING INSTRUCTIONAL READING LEVEL:
AN INVESTIGATION OF THE RELATIONSHIP AMONG
STANDARD CLOZE TESTS, MULTIPLE CHOICE CLOZE TESTS
AND THE INFORMAL READING INVENTORY

By

Susan Lubet Homan

December 1978

Chairman: William R. Powell
Major Department: Curriculum and Instruction

The relationship of a new form of cloze test, multiple choice cloze (MCC), to standard cloze and the Informal Reading Inventory (IRI) was explored in this study. The intent was to provide new information that would assist the classroom teacher in determining the instructional reading level of all students as easily and accurately as possible.

A secondary purpose of the study involved a comparison of MCC tests. The readability levels of the MCC passages were determined by traditional methods and by a new system of readability determination, Rasch calibration.

A third aspect of this study focused on placement decisions based on the two types of scoring criteria used with the IRI to determine instructional reading level.

Second, fourth, and sixth graders were participants in the study. Similar results at all three grade levels suggest generalizability of results for elementary and intermediate grade levels.

Correlations between standard cloze and MCC were low (ranging from $r = .27$ to $.80$), considering the same students were given the same stories in standard cloze and MCC forms. These results raise some question as to whether both forms are measuring the same type of reading comprehension.

High positive correlations were found between MCC passages with readability levels determined by traditional formula and MCC passages with readability levels based on Rasch calibration, indicating that these two methods of readability determination yield similar results.

A significant difference existed in placement of students by the Powell and Betts IRI criteria. These differences in placement indicate that classroom teachers should carefully choose the IRI scoring criteria they will use based on a conviction of accuracy of placement.

Three major implications for future research and practice were derived from this study. There is some evidence that MCC does not measure the same type of reading comprehension as measured by standard cloze or the IRI. The information on

this issue is inconclusive and further study is indicated before more specific conclusions can be reached. The relationship between standard cloze and the IRI also appears tenuous. The scoring criteria used for the IRI can make very significant differences in terms of accurate placement.

CHAPTER I INTRODUCTION

Every fall, at the start of the school year, teachers begin the process of assessing the knowledge and abilities of their new students. As part of this process the teacher seeks to determine the level at which each student can successfully read and assimilate information.

It is of vital importance that each student receive reading instruction at the proper level. That is, the books and materials they use in class should be at a readability level that is neither too easy nor too hard for each individual student. This specific level, at which materials are challenging for the student without being frustrating is called the instructional reading level. In order for each student to make optimum developmental reading progress, all materials and books should be at his/her instructional reading level (Dunkeld, 1970).

Classroom teachers continually face the problem of identifying the instructional reading level for every child in the class. The responses to this problem have been varied. The commonly used ways to determine instructional level are teacher judgement, standardized reading tests, Informal Reading Inventories (IRI), and cloze tests (Oliver, 1970).

Teacher judgement involves the placement of students in reading books based purely on the teacher's own subjectively conceived ideas. Teacher judgement has the weaknesses inherent in any measurement based totally on one person's opinion. It is very subjective and often inaccurate. Research indicates (Millsap, 1962) that teachers are unaware of frustration reading level among pupils in basal readers 30 percent of the time. Students are placed in basal readers at their instructional level only 70 percent of the time when teacher judgement is used. Although teacher judgement could be improved and stabilized with proper training (Millsap, 1962), this is rarely, if ever, done.

Standardized reading tests are group performance instruments. Patty (1969) concluded that the use of standardized grade equivalent scores is not a valid basis for determining instructional level. Standardized test norms appear to overestimate pupil instructional levels by at least one year (Killgallon, 1942; Patty, 1969; Sipay, 1961; Williams, 1963). Teachers who place students in basal readers on the basis of standardized test scores usually place them at their frustration reading level (Millsap, 1962).

The IRI is considered to be an accurate and often used method of determining instructional level. However, the IRI must be individually administered and is, therefore, time consuming for the classroom teacher. It takes approximately twenty minutes to properly administer an IRI to a student.

The standard cloze test also provides an accurate measure of instructional level. Cloze procedure as developed by Wilson Taylor (1953) was designed as an instrument for measuring the effectiveness of communication. Taylor described a functional unit of cloze measurement as a successful attempt to reproduce accurately a deleted part of a passage. The decision concerning the word to be reproduced was made from the remaining context of the passage. In effect, the reader must comprehend well enough to predict the word that is missing. The reader continuously draws from context clues to predict the nature of the language immediately ahead (Porter, 1976).

A cloze test is developed by taking a passage of approximately 250 words, leaving the first sentence intact, and then deleting every nth word. It can be administered to the entire class at one time and, therefore, is comparatively less time consuming than the IRI.

The cloze test, however, has not gained wide acceptance by classroom teachers. This is due to several factors. The cloze test is often very frustrating for the students. Also, all the blank spaces in a standard cloze passage are sometimes anxiety provoking. Students who miss 50 percent of the items feel they have failed or done poorly. However, according to all existing cloze criteria, 50 percent correct indicates the passage is at that student's instructional reading level (Alexander, 1968; Bormuth, 1967; Bormuth, 1968; Rankin & Culhane, 1969).

Teachers also object to standard cloze tests because of the scoring procedure. Only exact word replications are accepted as correct. Synonyms are rejected. Many teachers feel that a synonym still indicates comprehension and should be scored as a correct answer.

Standard cloze tests must also be hand scored which many teachers find cumbersome and time consuming. Some criticism of the standard cloze procedure centers around the argument that more than comprehension is necessary to complete a cloze unit. Even though a passage may be basically understood, a student might not be able to actually produce, or write in, the missing words. Especially in the case of students who speak in a dialect, or for whom English is a second language, standard cloze test scores might easily underestimate their true instructional level. It might be hard for those students to make an exact word replication even if they fully comprehend the passage (Porter, 1976).

Many of the aforementioned problems could be eliminated by a multiple choice cloze (MCC) test. Multiple choice cloze tests are based on the principles of standard cloze. The important difference between the two forms is that instead of leaving every nth word blank, three to five possible answers are given for each cloze unit. The student's task is to identify the correct word, rather than produce it. Since this involves circling or checking the correct word, rather than producing the right answer, it is much less frustrating to the student. The multiple choice format would also allow for much

simpler scoring with either machine scoring or the use of an overlay.

It currently takes teachers approximately three weeks to place students accurately in reading books. The student's enthusiasm for learning during the first three weeks of school is probably at its highest point. A more efficient method of determining instructional level would allow both teachers and students to capitalize on this eagerness to learn. Using MCC tests students could be placed in reading groups by the end of the first week of class, if not sooner.

Statement of the Problem

The primary purpose of this study was to investigate the relationship among standard cloze, MCC, and the IRI. If a close relationship among all three variables had existed, the second intention was to use standard cloze as a base to set up a viable criterion for MCC test scores in regard to instructional reading level. Multiple choice cloze tests could then be used by classroom teachers to determine the instructional reading levels of their students.

A third purpose was to investigate the validity of MCC tests as a measure of instructional reading level. The IRI was used as a base in determining this aspect of the validity of MCC tests.

Another consideration was a comparison of the IRI instructional level scores as scored by the Powell (1969, 1978) and Betts (1950) scoring criteria.

An additional variable was readability. Passages graded first through seventh grade level by traditional readability formulas were compared to a recent innovation in readability, Rasch calibration. Two forms of MCC tests were compared; the readability of one set was determined by traditional formulas (Dale-Chall, Harris-Jacobson, and Spache), and the other set's readability levels were Rasch calibrated.

Objectives and Research Questions

Listed below are the objectives of this study. Immediately following each objective are the specific questions that will be studied.

The first three questions are preliminary to the empirical validation of the instruments. They offer empirical evidence of certain underlying assumptions; namely, that MCC tests are different from standard cloze tests, that there is a difference in passage readability at each level, and that there is an interaction between different test forms and different levels.

Objective I

To investigate the similarity between standard cloze and MCC tests for the same passages, with the same students.

Question IA

Will there be differences between the mean scores for standard cloze and MCC passages for the same group of students?

Question IB

Will there be differences among the mean scores on the three levels of difficulty (readability levels) tested for each group?

Question IC

Will there be an interaction between the form used (standard cloze or MCC) and the level of difficulty (readability level) of the passage?

Objective II

To examine the strength of the relationship between standard cloze and MCC passages for the purpose of determining whether standard cloze scores can be used to predict MCC scores.

Question II

Will a correlation of .70 or more exist between standard cloze and MCC? The minimal criterion value of .70 was chosen for practical significance because this would indicate the shared variance of the two measures was close to 50 percent.

Objective III

To investigate the relationship among standard cloze scores, MCC scores, and the IRI.

Question IIIA

Will there be a relationship between standard cloze scores and IRI scores?

Question IIIB

Will there be a relationship between MCC scores and IRI scores?

Question IIIC

Will the correlations be greater between MCC scores and IRI instructional level scores or between standard cloze scores and IRI instructional level scores?

Objective IV

To determine the reliabilities of the standard cloze, MCC, and Kidder cloze passages.

Question IV

Will multiple choice cloze passage reliabilities be the same as standard cloze passage reliabilities at the same level?

Objective V

To study the relationship between two forms of MCC tests, Homan MCC and Kidder MCC.

Question V

Will there be a relationship between total Homan MCC scores and total Kidder cloze scores?

Objective VI

To investigate the relationship between IRI's instructional level scores as scored by the Powell criteria and Betts criteria.

Question VIA

Will there be a relationship between IRI's instructional level scores as scored by the Powell and Betts criteria?

Question VIB

Will there be a difference between instructional level means as scored by the Powell criteria and Betts criteria?

Question VIC

When the Powell and Betts criteria do not place students at the same instructional level, will a significantly greater proportion of students be placed at a higher instructional reading level by the Powell criteria than the Betts criteria?

Question VID

Will students be placed at the same instructional level by the Powell and Betts IRI criteria at least 75 percent of the time?

Definition of TermsBetts Criteria

The Betts criteria are the traditional standard set up by Emmett Betts (1950) for determining instructional reading level from an IRI. The Betts instructional level criteria are:

| <u>Book Level</u> | <u>Word Pronunciation</u> | <u>Comprehension</u> |
|-------------------|---------------------------|----------------------|
| All | 95% to 98% | 75% to 89% |

Frustration Level

Frustration level is the reading level at which the individual is "thwarted or baffled by the language (i.e., vocabulary, structure, sentence length) of the materials" (Betts, 1950, p. 152).

Independent Level

Independent level is the highest reading level at which the individual can read with full understanding and freedom from mechanical difficulties (Betts, 1950).

Informal Reading Inventory (IRI)

An IRI is a series of graded passages with comprehension questions for each passage. It is used to assess a student's level of reading (instructional, frustration, independent).

Instructional Level

Instructional level is the highest reading level at which systematic instruction can be initiated (Betts, 1950). It is the reading level at which the student is challenged by the material without being frustrated.

Multiple Choice Cloze (MCC)

Multiple choice cloze tests are constructed as standard cloze tests and based on the same underlying principles. Every nth word is deleted from a passage of approximately 250 words. Each blank space or item contains the correct answer plus two to four distractors.

Multiple choice cloze can be easily scored by machine or by hand using an overlay. The MCC test retains the best points of standard cloze tests while reducing student anxiety and enhancing ease of scoring. Evidence also seems to indicate it is a more valid measure of reading comprehension than standard cloze tests (O'Reilly & Streeter, 1977; Porter, 1976).

Powell Criteria

The Powell criteria are the differentiated criteria set up by William Powell (1969, 1978) for determining instructional reading level from an IRI:

| <u>Book Level</u> | <u>Word Pronunciation</u> | <u>Comprehension</u> |
|-------------------|---------------------------|----------------------|
| PP-2 | 87% to 94% | 55% to 80% |
| 3-5 | 92% to 96% | 60% to 85% |
| 6+ | 94% to 97% | 65% to 90% |

Semantically Correct Word

A semantically correct word is similar in meaning to the deleted word in a functional cloze unit. It would make sense in context if substituted for the deleted word.

Example. The horse _____ over the tree. Correct answer: JUMPED; semantically correct answer: FELL.

Standard Cloze Test

A cloze passage is any passage which has every nth word deleted. No deletions are made in the first one or two sentences to allow the theme of the passage to be established. The reader's task is to fill in the blanks with an exact replication of the missing word. Cloze tests are distinguished from completion tests by the fact that cloze test deletions are made using a set of mechanically objective and prespecified rules, while the deletions in completion tests may be made using subjective concepts.

Syntactically Correct Word

A syntactically correct word is the appropriate part of speech (noun, verb, function, modifier) needed to complete a functional cloze unit. It is the same part of speech as the correct word, but does not necessarily make sense if substituted for the correct word.

Example. The _____ jumped over the tree. Correct answer: HORSE: syntactically correct answer: CHAIR.

Limitations of the Study

There are five points to be considered as possible limitations of this study. The sample used was not randomly selected from the total population. Therefore, generalizing from these results should be done with caution.

The second graders were tested in only two sessions. Due to the age, attention span, and slow rate of speed at which second graders work, many of them tired and/or gave up before they had completed all the passages. It took many second graders an hour to complete the passages when the standard cloze and Kidder cloze passages were administered together.

Three limitations involved the IRI. The comprehension questions for the IRI were limited in number, only five per passage. With only five questions, a student who missed one question would automatically be at instructional level. The student might actually be able to handle the material independently, but due to carelessness, or misunderstanding, one question was missed. Five questions allow for little differentiation between independent, instructional, and frustration reading levels.

The students were able to reread the IRI to locate specific information. Several of the comprehension questions directed the student to "look in the passage for," or "find"

a specific word. This allowed for some rereading of the passage which might affect the results. However, the rereading was only to locate specific words or information.

In addition, IRI's were given to only thirty to thirty-three students at each grade level. Giving every student an IRI would have presented more complete data for analysis. However, due to the time involved in administering IRI's to 548 students, it was not feasible for this study.

CHAPTER II REVIEW OF RELATED LITERATURE

History and Development

Wilson Taylor (1953) is credited with the development of cloze procedure. He introduced it as, ". . . a new psychological tool for measuring the effectiveness of communication" (Taylor, 1953, p. 123). He recognized and tested its usefulness as a new approach to readability. Cloze procedure was seen as ". . . a method of intercepting a message from a 'transmitter' (writer or speaker), mutilating its language patterns by deleting parts, and so administering it to receivers (readers and listeners) that their attempts to make the patterns whole again potentially yield a considerable number of cloze units" (Taylor, 1953, p. 126).

An actual cloze unit was considered any single occurrence of a successful attempt to reproduce accurately a deleted part of a passage or message. The method used to supply the missing part was determined from the remaining context (Taylor, 1953).

The term "cloze" comes from the Gestalt concept of closure (Taylor, 1957). This concept relates to the tendency to "see" a not quite complete circle as a whole circle by mentally completing the picture. The same concept, Taylor assumed, held true of people trying to complete a mutilated

sentence by filling in the words that made the finished pattern of language symbols fit the apparent meaning.

Ohmacht, Weaver and Kohler (1970), in an attempt to factor analyze cloze procedure, found only a moderate correlation between cloze factors and perceptual cloze factors. The actual relationship between the Gestalt concept of closure is still open to question.

Another aspect of cloze procedure investigated by Taylor (1953) was its use for determining readability. In a correlational study using the Dale-Chall (Dale & Chall, 1948) and Flesch (Flesch, 1948) readability formulas, cloze procedure was shown to be an effective and reliable method of contrasting the readability levels of various passages.

Taylor (1957) assumed that readability and comprehensibility were essentially synonymous terms. He saw the readability level as being the same as the understandability of a passage. This would seem to be true if readability formulas are accurate. Generally, the lower the readability level of a passage, the easier it is to understand.

Taylor (1953) specified that cloze procedure does not deal directly with specific meanings of words. The total context and the redundancy of our language are actually involved. Cloze procedure counts the instances of language usage correspondence, and not the meanings themselves. This relates to a measure of the likeness between a writer's patterns and the patterns the reader is anticipating while reading. Agreement between the writer's patterns and the reader's anticipated

patterns create a language usage correspondence. Taylor (1953) credited ideas from the total language context concept, Osgood's dispositional mechanisms, and statistical random sampling as contributing to the development of cloze procedure.

In simplistic terms, cloze procedure involves leaving every nth word of a passage blank. The student then fills in those blanks using the context to determine the missing word.

Cloze procedure has been and continues to be studied. Taylor (1953) saw its potential as an approach to readability. Believing that readability and comprehensibility were similar, if not the same, his original studies explored heavily in that area. Taylor developed the form of standard cloze most often used today.

Methodological Considerations

Any passage may be used as the basis for a cloze test. In his original study, Taylor (1953) experimented with two types of deletions, random, and every nth word. Random deletions were chosen by a random number generator in accordance with the belief that if enough words were struck out at random, the blanks would come to represent proportionately all kinds of words to the extent that they occur in the passage. The second and more convenient method was every nth word deletions.

Several studies have supported the indication that every fifth word deletion discriminates best in passages above third grade level (Rankin, 1959; Rankin, 1965; Taylor, 1953), unless the passage is extremely technical in which case every tenth

word deletions are recommended (Cranney, 1968; Taylor, 1953). In a separate study, MacGinitie (1961) found that additional context beyond five words did not help in the restoration of missing words. He also concluded that omitting every third word made restoration difficult.

Salzinger, Portnoy and Feldman (1962) concluded that having six words on either side of the cloze blank did not produce more correct guesses than leaving four words on either side of the cloze item. They felt students were either unable to, or simply did not, make use of context of more than five words on either side of the cloze blank.

Fillenbaum, Jones and Rappoport (1964) did a study using deletions every two words up to deletions every six words. They found the greatest differences in performance between the passages with deletions every two words and deletions every three words. They pointed out that with deletions every two words the students were sometimes able to guess form, but rarely capable of replacing the words correctly.

The length of Taylor's (1953) original passages was approximately 175 words, allowing for thirty-five deletions. Taylor concluded that scores tended to stabilize after the first twenty to forty words of an any-word type of deletion passage. In a later study, Taylor (1956) recommended fifty deletions as a suitable length for cloze tests. No statistical support was given for the fifty deletion suggestion.

Rufener (1972) used shorter passage lengths in deference to the young age of her subjects. The number of deletions

ranged from eighteen for the second grade level passage to twenty for the sixth grade level passage. Rufener used cumulative scores based on percentages to ascertain the stability and reliability of cloze scores. She found stable scores on some passages within the range of ten to twenty deletions, or fifty to one hundred words in the passage. She concluded that stability of scores was more a function of the individual cloze passage than the length of the passage.

A final question concerning cloze test construction involves rational versus mechanical deletion of words. Mechanical, also called any word deletion, refers to every n th word deletion regardless of the grammatical form of the word. Rational deletion involves deleting only words that have a special grammatical function (Rankin, 1959). These words may be only nouns, verbs, adjectives, and adverbs, as is the case for lexical cloze. Lexical cloze involves the meaning of individual words. This differs from what standard cloze is assumed to actually measure, structural meaning. Structural meaning is signaled by a system of morphological and syntactical clues apart from words as vocabulary units.

Taylor (1953) indicated that if only important words such as nouns, verbs, adverbs, and adjectives were deleted it might not reflect the passage meaning accurately. Two passages of equal length might contain a very different number of meaning conveying words, possibly as much as double the amount in the passage containing the fewest nouns, verbs, adverbs,

and adjectives. The effect of this on the subjects should be included in the results of the cloze test.

Rankin (1965) discovered that mechanical deletion produced a sizeable number of nondiscriminating items which might lower the reliability of the cloze test. However, the number of items (deletions) in the passage could be increased, thereby increasing the reliability of the passage. Rankin also indicated that any word deletion cloze tests correlated more highly with the criterion tests of prereading knowledge, recall, and aptitude. Due to this he considered mechanical deletions superior.

Greene (1965) compared mechanical and rational deletions. The rational deletions produced higher reliabilities and item discriminations. However, Greene noted that the difference in time necessary to construct the rational cloze tests often outweighed the advantage of a slight gain in reliability, especially since there were no significant differences between the mean scores of the students.

Bloomer (1966) and Louthan (1965) both investigated various forms of rational deletions. Bloomer (1966) concluded that marked comprehension losses occur "where the structure of the prose and its meaning are broken by deletions of the basic meaning carriers of the language, i.e., nouns, verbs, and modifiers" (p. 66). His results suggest that the part of speech deleted does have an effect on the ability of the individual to complete cloze procedure tests.

Louthan (1965) indicated that systematic deletions of structural function words (nouns, verbs, adjectives) would produce a loss in reading comprehension, regardless of the number of deletions. In a similar study, Rankin (1959) reported that mechanical deletions correlated significantly higher with reading comprehension sections of the Diagnostic Reading Test than rational deletions.

In scoring cloze tests, points may be awarded for exact word only or synonyms accepted. Taylor (1953, 1956) recommended giving credit for exact word answers only since this retains the objectivity of cloze procedure. Accepting synonyms has been suggested as beneficial when using cloze procedure as a teaching technique (Jongsma, 1971). However, if synonyms are counted as correct, the subjectivity of the scorer becomes an issue. Also, Cronbach, as paraphrased by Taylor (1953) states that cloze tests scored objectively satisfy the assumptions for true scores and can be considered as such.

The areas covered by methodological considerations all involve deletions. The literature reviewed indicates a strong preference for using every fifth word deletion. The number of deletions has been known to range from eighteen to over fifty and still maintain relatively high and stable reliabilities.

The issue of mechanical versus rational deletion has proponents on both sides. Most of the researchers favor mechanical deletion even though that method often lowers passage reliabilities. There is almost total agreement on scoring by exact answer only. Once the subjectivity of the

scorer becomes a factor, the cloze scores could no longer be considered true scores. Many of these issues are still under investigation (O'Reilly & Streeter, 1977).

Reliability

Taylor's (1953) original study produced a cloze test internal consistency measure of .56 using Kendall's W. Rankin (1965) pointed out that mechanical selection of words might produce a sizeable number of nondiscriminating items which lower reliability. However, other studies of cloze procedure have yielded very high reliabilities. In a validity study involving cloze procedure, Bormuth (1969) used split-half reliabilities and produced highly satisfactory reliability coefficients ranging from .92 to .94. In a study involving mechanical and rational deletion of cloze procedure, Greene (1965) found K-R #21 reliabilities of .76 and .90. Cranney (1973), in a study using cloze in both standard and MCC tests, found reliabilities using the K-R #20 formula which ranged from .83 to .93.

Using MCC cloze tests, O'Reilly and Streeter (1977) found very high estimates of internal consistency. The K-R #20 reliabilities ranged from .94 to .97 with a median reliability of .96.

The cloze reliability data available strongly indicated that standard cloze and MCC tests were reliable instruments. However, no overt comparisons were made between standard and MCC reliabilities.

Criteria

Bormuth (1967) recognized that a frame of reference by which to interpret cloze scores was lacking. In a landmark study, Bormuth correlated cloze test scores with multiple choice scores to produce workable criteria. In this study Bormuth used one hundred fourth and fifth graders. The standard cloze test had every fifth word deleted and had fifty items. By regressing cloze scores on multiple choice scores, a 38 percent correct standard cloze score was equated with a 75 percent correct multiple choice score. The 38 percent was the lower limit for instructional reading level. A standard cloze score of 50 percent was comparable to 90 percent correct on the multiple choice test. The standard error of the estimate, however, was six percentage points.

In a later study, Bormuth (1968) correlated cloze scores with word recognition and comprehension. The Gray Oral Reading Test was used. The cloze tests contained fifty items per test and had every fifth word deleted. A total of 120 fourth, fifth, and sixth graders comprised the sample, forty at each grade level. A matching procedure was used to determine comparable standard cloze and comprehension scores. The results indicated cloze scores of 44 percent were comparable to comprehension scores of 75 percent. The upper range equated cloze scores of 57 percent to comprehension scores of 95 percent. The cloze scores comparable to word recognition scores of 95 percent and 98 percent were 33 percent and 54 percent,

respectively. It should be noted that both of these studies are based on relatively small samples, especially considering the generalizability of these scoring criteria.

Using a sample of 105 fifth graders, Rankin and Culhane (1969) replicated Bormuth's study and obtained similar results. Fifty item cloze tests with every fifth word deleted were administered to the students along with multiple choice tests from the same articles. Regression equations were set up to predict multiple choice percentage scores. The criteria that evolved equated 41 percent correct on a cloze test to 75 percent correct on the multiple choice test. Also, 61 percent correct on cloze tests was comparable to 90 percent on the multiple choice tests. While Rankin and Culhane recommend that teachers interpret cloze scores by the criteria from this study "with some degree of confidence" (Rankin & Culhane, 1969, p. 197), it should be noted that this study involved only one grade level.

In a separate study, Alexander (1968) set up criteria for cloze scores based on instructional level IRI scores. Alexander used data from 365 students in his sample from grades four, five, and six. The number of cloze items ranged from forty-five to ninety-two. Every fifth word was deleted. Alexander's criteria had an instructional level range of 47 percent to 60 percent based on the IRI's instructional level scores. The IRI instructional level scores were determined using the Powell criteria (Powell, 1969).

All four of these studies are based on fourth only, fifth only, or fourth, fifth and sixth grade students. There are possible differences among these grade levels that discourage direct comparison of these cloze criteria.

These standards for determining cloze instructional level are all based on varying base criteria. However, the 75 percent to 90 percent correct seems constant, even though what the 75 percent to 90 percent measure changes with each study, covering the range of multiple choice tests, oral reading tests, comprehension questioning and the IRI.

Measurement of Reading Comprehension

Conflicting evidence exists in the literature concerning the validity of standard cloze tests as a measure of comprehension. Bormuth (1969) in a factor analytic study of cloze tests found that one factor accounted for 77 percent of the variation in the correlation matrix. Using a principal components analysis, Bormuth analyzed correlations among nine cloze tests and seven multiple choice comprehension tests. He concluded ". . . cloze made by deleting every fifth word measure skills closely related or identical to those measured by conventional multiple choice reading comprehension tests" (Bormuth, 1969, p. 365).

Porter (1976) stated that "the fact that the ability to predict what lies ahead depends on the ability fully to comprehend the language being processed at any given moment

provides the justification for cloze procedure as a test of comprehension" (p. 152).

In another factor analytic study, Horton (1974-75) used cloze tests from science and social studies, paragraph reading tests, and twelve tests designed to measure structure of intellect. He performed several analyses and concluded that the factors appeared to be invariant from one analysis to the others. The study attempted to establish both construct and concurrent validity for cloze tests. The construct was defined as the ability to deal with the linguistic structure of the language as related to the ability to deal with relationships among words and ideas. Horton suggested the variance shared among cloze tests and reading comprehension tests, reading gain tests, and verbal intelligence tests are probably a measure of the student's ability to deal with relationships among words and ideas. The correlations established concurrent validity.

Weaver and Kingston (1963) also factor analyzed cloze tests. College level students participated in the study. The tests used were the Davis Reading Test, several subtests of the Modern Language Aptitude Test, the STEP Listening Test, the Ohio State Psychological Examination, and eight cloze tests. Weaver and Kingston concluded that cloze procedure was only moderately related to verbal comprehension. Much specific variance was unexplained by any of the factors.

Bormuth (1969) questioned the interpretation of Weaver and Kingston's (1963) data for four different reasons. The subjects were college students which produced a less varied ability distribution than most elementary school level studies. The correlations on which they based their calculations were different from those obtained from other investigators. The standardized tests used had unusual factor loadings, and the cloze tests had inconsistencies in loading patterns among themselves.

Carroll (1972) in a discussion of comprehension and various tests of comprehension stated, "Cloze technique in its usual form is too crude to permit measuring the degree to which the individual comprehends particular lexical or grammatical cues, or possesses a knowledge of specified linguistic rules" (p. 19). He suggests there is no clear evidence that cloze scores measure the ability to comprehend major ideas or concepts that run through a discourse.

The evidence cited by Carroll as leading to this conclusion is the Weaver and Kingston factorial study (1963) previously discussed, a study by Coleman and Miller (1968), and Rankin's (1958) study. The Coleman and Miller (1968) study, as depicted by Carroll (1972), attempted to measure knowledge gain from a cloze passage. A distinction between knowledge gain and comprehension or understanding is arguable. Rankin (1959) concluded scores based on the deletions of nouns and verbs seem to measure something different than cloze scores

based on deletion of function words. This conclusion has no direct implication in terms of the ability of a mechanically deleted cloze passage to measure reading comprehension. Therefore, Carroll (1972) does not seem to have strong or clear supportive evidence for his statements.

In another study using college students (Anderson, 1974) different conclusions were reached. Anderson found cloze tests were valid measures of comprehension. Anderson also noted that there were more correct responses on the cloze tests when a MCC format was utilized.

Porter (1976) expressed his view that standard cloze tests do not measure what they claim to measure. Although comprehension of the passage is necessary to the successful completion of a standard cloze passage, that alone is not enough. A student being administered a standard cloze test must produce language to fill in each cloze space. Porter stated "to be valid, a test should measure what it is intended to measure, and language production is not comprehension, neither are the two necessarily concomitant" (p. 152). A student may understand or comprehend a cloze passage and still be unable to predict the correct word for a cloze item. This is especially relevant for students who speak in a dialect, or for whom English is a second language.

Porter (1976) proposed devising a method of testing comprehension with the advantages of standard cloze procedure, but without necessitating language production. The solution he suggested was a MCC test. Although MCC would take a little

longer to construct due to the added distractors, the added test validity would be worth the additional time and effort. Also, the process of selecting alternatives offers the possibility of more control and flexibility than standard cloze procedure possessed. This control is exemplified by being able to choose distractors which vary according to the "depth of linguistic attainment and fineness of stylistic discrimination of the student" (Porter, 1976, p. 154).

An additional virtue of MCC is that it can be constructed so that an overlay can be easily used for scoring, or it can be used with answer sheets and machine scored. O'Reilly and Streeter (1977) also point out that a MCC test has greater face validity as a measure of comprehension.

Guthrie (1973) correlated MCC tests with the Gates MacGinitie comprehension and vocabulary test to measure the validity of MCC as a test of reading comprehension. The MCC test correlated .85 with the vocabulary section of the Gates MacGinitie and .82 with the comprehension section. The subjects in this study were thirty-six elementary aged normal and disabled readers.

Cranney (1973) demonstrated that a MCC test when correlated with the Cooperative Reading Test had a validity coefficient that showed 25 percent shared variance with the comprehension sections of the Cooperative Reading Test.

A recent study of O'Reilly and Streeter (1977) stated the belief that multiple choice technique would retain many

of the good points of standard cloze procedure, such as absence of questions, and objective item construction, while improving its applicability as a measure of reading comprehension. Also, due to the format of MCC, students would suffer less from test anxiety which should increase the reliability of the tests. The excessive difficulty and ambiguity of the standard cloze test would be greatly reduced.

O'Reilly and Streeter (1977) conducted a factor analytic study of MCC tests using tests based on Bormuth's Wh-Items format, the California Achievement Test, a Test-Wiseness Test, and the Short Form Test of Academic Aptitude. They concluded that MCC tests were a measure of reading comprehension that was essentially independent of IQ.

The main factor of the MCC tests appeared to be one of literal comprehension. However, there was evidence of two other factors possibly representing other forms of comprehension.

Summary of Research

Cloze procedure is backed by a proliferation of research as a tool for measuring some sort of reading comprehension. A dialogue continues concerning the actual type and amount of comprehension it measures.

Many researchers have attempted to establish criteria by which to determine instructional reading level from cloze tests. All of the research in that area and in most other areas of cloze research has been limited to fourth grade

students and above. The field is open for investigating primary and intermediate grade children to see if results are generalizable across grade levels.

The newest area of research in cloze procedure appears to be in the area of MCC. The relationship between this new form of cloze and standard cloze is yet to be studied.

Information on the relationship of standard cloze scores and the IRI is also very limited. Exploration of the relationships and interrelationships in the aforementioned areas of MCC, standard cloze, and the IRI might well be of great potential value to the classroom teacher. Through research in these areas, a way to determine instructional reading level with ease, accuracy, and reliability might yet be forthcoming.

CHAPTER III PROCEDURES AND METHODOLOGY

Sample

The subjects were 557 students selected from a larger school population of second, fourth, and sixth grade pupils. Three schools from the Pinellas County School System participated in the study: Dunedin Elementary School, Lealman Elementary School, and Mt. Vernon Elementary School. The total second, fourth, and sixth grade population of Pinellas County is 18,729.

These three schools were selected because they housed first through sixth grades and the principals and faculties were willing to help with the study. The schools were representative of three different socioeconomic levels. This determination was made according to data on free and reduced lunches. The students who receive free and reduced lunches are determined by amount of family income. Therefore, classification on this basis is justified.

Lealman Elementary School had 78 percent of its students on the free or reduced school lunch program. This was considered the low socioeconomic school. Mt. Vernon Elementary School had 47 percent of its students on the free or reduced school lunch program. Mt. Vernon was considered the middle

socioeconomic school. Dunedin Elementary School had 19 percent of its students participating in the free or reduced lunch program, and was considered the high socioeconomic school.

The students were not assigned to classes on the basis of reading ability. All students in the class were tested, including those involved in special programs such as Special Learning Disabilities and Emotionally Handicapped. Absentees were not given makeup tests. Students with incomplete test data were eliminated from the study.

Instruments Used

Each student was tested on nine passages. The investigator wrote original passages ranging in readability levels from 1.5 to approximately 7.5. Three of these passages were presented to the student as standard cloze passages. The second grade students received first, second, and third grade level passages. Fourth grade students were required to complete passages for grade levels three, four, and five. Sixth grade students were tested with fifth, sixth, and seventh grade level passages.

The students were also tested with the same three passages or stories in an MCC format (Homan Cloze).

Additional MCC passages (Kidder Cloze) written for the New York School System and Rasch calibrated for readability were also given to each student. Three of these passages (also covering three grade levels) were completed by each student. This was done for two purposes: it served as a

check on the readability level of the original passages; and it checked on the possibility of an effect on the Homan Cloze tests caused by the student's previously taking a standard cloze test on the same story.

Thirty-six students were randomly selected at the second grade level, thirty-four at the fourth grade level, and thirty-six at the sixth grade level, to be administered IRI's. This served as a base measure of instructional reading level. It was used to establish the validity of the Homan Cloze and to revalidate standard cloze as a measure of instructional reading level.

Instrument Development - Original Passage

Purpose of Instrument

Stories of approximately 135 words were written at first through seventh grade reading levels. These passages were made into standard cloze and MCC tests to be used to determine the instructional reading levels of second, fourth, and sixth graders. Both standard cloze and MCC are considered measures of reading comprehension which is a direct reflection of reading level. If a student cannot comprehend material sufficiently at a certain level, he/she cannot effectively work and learn at that level of instruction. The original cloze passages, both standard and multiple choice (Homan Cloze) were intended to provide a measure of reading comprehension that could be used to determine instructional reading level. These passages are included in Appendices B and C.

Source of Items

Every fifth word was deleted from the cloze passages. Taylor (1953, 1956) and Rankin (1965) both have supported fifth word deletion. MacGinitie (1961) found that additional context beyond five words did not help in restoration. In a later study, Salzinger, Portnoy and Feldman (1962) also determined that having six words on either side of the cloze blank did not produce more correct guesses than having four words on either side of the blank. They concluded that subjects either cannot or do not utilize context of more than five words on either side of a blank. An additional conclusion was that words predicted for every fifth blank are independent of each other.

The following list contains the number of words, number of deletions, and readability levels of the original standard cloze passages. The Dale-Chall readabilities are approximate grade levels derived from Dale-Chall Formula scores (Dale & Chall, 1948).

| | Words | Deletions | Readability | Source |
|-----------|-------|-----------|-------------|-------------------------------------------|
| 1st grade | 100 | 18 | 1.5 | Spache, 1977 |
| 2nd grade | 138 | 23 | 2.5 | Spache, 1977 |
| 3rd grade | 126 | 23 | 3.5 | Spache, 1977 |
| 4th grade | 147 | 23 | 4.8 | Harris-Jacobson |
| 5th grade | 126 | 23 | 5.5 | Dale-Chall (converted from Formula Score) |
| 6th grade | 133 | 23 | 6.5 | Dale-Chall (converted from Formula Score) |
| 7th grade | 124 | 21 | 7.5 | Dale-Chall (converted from Formula Score) |

The following is a list of the number of words, deletions, and readability levels of the Kidder Cloze passages:

| | Words | Deletions | Readability | Source |
|-----------|-------|-----------|-------------|------------------------------------------------------------------|
| 1st grade | 113 | 18 | 1.4 to 1.7 | Rasch calibrated and Spache Formula for first, second, and third |
| 2nd grade | 88 | 13 | 2.5 | |
| 3rd grade | 66 | 10 | 3.5 | |
| 4th grade | 65 | 10 | 4.66 | Rasch calibrated and Dale-Chall Formula |
| 5th grade | 61 | 10 | 5.42 | Score for fourth, fifth, sixth, and seventh |
| 6th grade | 74 | 11 | 5.84 | |
| 7th grade | 70 | 12 | 6.38 | |

The MCC tests at all levels had three selections for each deleted word. One choice was the correct answer. One of the selections was syntactically correct but semantically incorrect. The third choice was neither semantically nor syntactically correct.

Four parts of speech were identified for syntactic agreement or nonagreement (Guthrie, 1973). These categories were noun, verb, modifier, and function. The noun group included nouns and pronouns. The verb group included verbs and auxiliary verbs. The modifier group included adjectives and adverbs, and the function group included prepositions, articles, and conjunctions.

The syntactically correct distractors for the first, second, and third grade passages were selected using systematic sampling from the Dale-Chall List of 769 words, and from the Dale-Chall List of 3000 words for the fourth, fifth, sixth,

and seventh grade passages (Dale & Chall, 1948). In each case the number of items in the cloze test was divided into the number of words on the list. The word list was then divided into equal groups using the derived quotient. If the first word of the group was not the needed part of speech, the next word on the list that was the correct part of speech was used. By using the words on the Dale-Chall lists the readability level of the distractors was controlled.

The choice that was neither syntactically nor semantically correct was selected from the passage using a random number table. The position of the correct answer was randomly assigned.

Scoring

In the standard cloze passages only the exact word was scored as correct. Synonyms were not allowed. Spelling mistakes were permitted as long as the word was still recognizable. By scoring only the exact word as correct, the objectivity of the scoring procedure was greatly enhanced.

The MCC test items were scored as right if the correct word was circled. If the correct answer and an incorrect answer were circled on the same item, it was considered a wrong answer.

Item Analysis

An item analysis was performed using biserial correlations between individual item scores and total scores for the passage. A biserial correlation is used for computing the correlation between one continuous variable and one artificially dichotomized

variable. The item scores were considered an artificial dichotomy since other words, or answers, could be inserted to make sense in the passage and so are "correct" in some sense.

No items were deleted based on this item analysis. However, this information may be of value for future research.

Reliability

Internal consistency estimates were computed as measures of reliability on all passages, standard and MCC.

Validity

Standard cloze has been shown to be a valid measure of reading comprehension (Bormuth, 1969; Rankin, 1959). Criterion data were collected by administering Information Reading Inventories to 107 randomly selected students. Multiple choice cloze and standard cloze scores were correlated with IRI instructional levels to determine their validities.

Other Instruments

Some MCC passages (Kidder Cloze) taken from the Test Development Notebook developed by the Bureau of School and Cultural Research, Division of Research, New York State Education Department were also administered to each student. The readability level of these passages had been determined in two ways. A conventional method, the Spache Readability Formula (Spache, 1953) was used for the first, second, and third grade passages, and the Dale-Chall Readability Formula (Dale & Chall, 1948) was used for the fourth through seventh

grade passages. Also, these passages, along with a series of MCC passages ranging from first through twelfth grade level, had their readability level determined by a Rasch calibrated scale. The Rasch based scale is considered to be a "person-free" estimate of passage difficulties (Kidder, 1977). That is, calibrations derived from the Rasch model are not sample dependent. For further information on the Rasch model, see Hambleton and Cook (1977), or Wright and Panchapakesan (1969).

These multiple choice passages served two purposes. They provided a check on the readability level of the original passages and were used to determine whether any learning took place from the standard cloze passages which were always administered first. Each student took three standard cloze tests and three Homan Cloze (MCC) tests on the same stories. They also took three MCC tests (Kidder Cloze) on different stories.

If learning had occurred from the standard cloze tests, the students would be expected to score higher on Homan MCC tests than Kidder MCC tests. The standard cloze tests and the Homan MCC tests were identical in content, differing only in form. Therefore, learning from the standard cloze test would very likely be reflected in a high Homan MCC test score.

Source of Items

The Kidder Cloze passages were taken as complete stories. The investigator then randomly deleted every fifth word using the same procedure used for the original MCC passages. The

same procedure was also followed for selecting the two distractors.

Reliability

Some of the Kidder Cloze passages had only ten items. Rufener (1972) achieved stable scores after only ten to twenty cloze deletions. Using cumulative percentage of successful scores, most passages indicated stability of scores after ten to fifteen deletions.

Using internal consistency as a measure of reliability, the Kidder Cloze passages had respectable reliabilities, ranging from .65 to .90 (see Tables 14, 15, 16). Based on the high reliabilities of all Kidder Cloze passages and the data from the Rufener (1972) study, it seems reasonable to conclude that even ten cloze items can yield reliable results. The three passage scores for each student were combined for correlation with Homan MCC. The combined passage reliabilities were higher than the individual passage reliabilities.

Informal Reading Inventory (IRI)

The IRI is a measure of a student's instructional reading level. Using graded passages, a measurement is obtained of comprehension and word recognition scores.

The IRI used in this study was developed by Alexander (1968). It consists of a series of graded stories, each followed by five comprehension questions (see Appendix E). The comprehension questions combined questions at the literal, inferential, vocabulary, and evaluative comprehension levels.

Alexander (1968) developed a cloze criteria for determining instructional reading level using this IRI.

The Powell criteria for determining instructional level from IRI scores were used. However, the scores were also evaluated using the Betts criteria. Differences when using each criteria were examined.

Thirty-six second graders, thirty-four fourth graders, and thirty-six sixth graders were chosen randomly and administered an IRI. Informal Reading Inventory scores were used as the criterion to validate MCC as a means of determining instructional reading level.

Procedure

Each class at all levels had two testing sessions of approximately thirty minutes each. The first session always involved three standard cloze passages. In half of the classes the Kidder MCC passages were also administered at the first session.

The second testing session, which was always at least one week after the first session, involved the administration of the Homan MCC test. Half of the classes were administered the Kidder MCC at the second testing session. The IRI's were given individually after both testing sessions were completed to the reduced sample (n=107).

Second grade students were administered standard cloze, Homan MCC, and Kidder MCC passages at the first, second, and third grade readability levels. Fourth grade students were administered standard cloze, Homan MCC, and Kidder MCC passages

at the third, fourth, and fifth grade readability levels. Sixth grade students were administered standard cloze, Homan MCC, and Kidder MCC passages at the fifth, sixth, and seventh grade readability levels.

Administration

All tests were administered by the investigator to insure uniformity of instructions. The instructions were given orally. A standard cloze item was demonstrated on the chalkboard. The example for second grade was: THE DOG RAN TO _____ DOOR. HE WANTED TO _____ IN. The fourth and sixth grade example was: THE SWING WENT BACK _____ FORTH.

There was no time limit on the tests. The students were permitted to ask their teacher or the investigator to spell words, but words could not be pronounced for them. (For full directions, see Appendix A.)

Method of Analysis

All data were punched onto IBM cards for subsequent analysis. The Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner & Brent, 1970) was used to facilitate computation. The Biomedical Computer Programs, P-Series (Dixon & Brown, 1977) was used for the analysis of variance computation. The research hypotheses now follow stated in null form. After each hypothesis is a brief explanation of the statistical analysis utilized.

Hypothesis IA

There will be no differences between the mean scores for standard cloze and MCC passages for the same group of students.

Hypothesis IB

There will be no differences among the mean scores on the three levels of difficulty tested for each group.

Hypothesis IC

There will be no interaction between the form used (standard cloze or MCC) and the level of difficulty (readability level) of the passage.

Analysis of variance was applied to compare the standard cloze scores and Homan MCC scores for the same students using a randomized block factorial design with the subjects serving as blocks. The students' scores on the three standard cloze tests were compared to their scores on MCC tests.

Differences due to form (standard cloze and MCC), and level of difficulty were tested. The degree of interaction between form and level was examined.

These hypotheses were tested at a significance level of $\alpha = .05$. Since the assumption of symmetry for the variance-covariance matrix was not tested, a conservative F test, the Geisser-Greenhouse test (Kirk, 1968), was used to test significance for both main effects and the interactions at the second, fourth, and sixth grade level analyses.

Hypothesis II

A correlation of .70 or more exists between standard cloze and MCC.

The strength of the relationship between standard cloze scores and MCC scores was assessed using the Pearson Product-Moment correlation. A preset correlational value of .70 was used to determine the practical significance of the relationship.

Hypothesis IIIA

There will be no relationship between standard cloze scores and IRI scores.

Hypothesis IIIB

There will be no relationship between MCC scores and IRI scores.

The Pearson Product-Moment correlation was computed to assess the relationship between both standard cloze and MCC to the IRI.

These hypotheses were tested at a significance level of $\alpha=.01$. The $\alpha=.01$ level of significance was chosen because the results of these tests may determine if the instruments involved will be used for placement. When considering placement of students for instructional reading level, accuracy is extremely important. The significance level of $\alpha=.01$ was used to guard against a Type I error. All hypotheses that relate to possible placement decisions were tested at a significance level of $\alpha=.01$.

Hypothesis IIIC

The correlations between MCC scores and IRI instructional level scores will be the same as the correlations between standard cloze and IRI instructional level scores.

The difference between the Pearson Product-Moment correlation coefficients of standard cloze and MCC to the IRI was tested for significance using the following formula:

$$t = (r_{12} - r_{13}) \sqrt{\frac{(N-3) (1 - r_{23}^2)}{2 (1 - r_{23}^2 - r_{12}^2 - r_{13}^2 + 2r_{23}r_{12}r_{13})}}$$

where,

r_{12} is the correlation coefficient of the IRI to MCC.

r_{13} is the correlation coefficient of the IRI to SC.

r_{23} is the correlation coefficient of SC to MCC

d.f. = N-3.

This formula (Guilford & Fruchter, 1973) is used to test for significance between correlation coefficients from the same sample.

Hypothesis IV

Multiple choice cloze passage reliabilities will be the same as standard cloze passage reliabilities at the same grade level.

The internal consistencies of all instruments were estimated using the Gitap Program (Baker, 1970). Hoyt reliability coefficients are computed by this program. Item difficulties and biserial correlations were also provided by this program.

Hypothesis V

There will be no relationship between total Homan MCC scores and total Kidder Cloze scores. Total cloze scores are the students' combined scores on all three Homan Cloze or Kidder Cloze passages.

The relationship between Homan MCC scores and Kidder MCC scores was assessed by the Pearson Product-Moment correlation method. This hypothesis was tested at a significance level of $\alpha=.01$

Hypothesis VIA

There will be no relationship between IRI's instructional level scores as scored by the Powell and Betts criteria.

The relationship between instructional level scores from the IRI as scored by the Powell criteria and the Betts criteria was assessed by computing the Pearson Product-Moment correlation. The hypothesis was tested at a significance level of $\alpha=.01$.

Hypothesis VIB

There will be no difference between instructional level means as scored by the Powell and Betts criteria.

The difference between instructional level means as determined by the Powell and Betts criteria was tested by a t-test for comparison of means. The hypothesis was tested at a significance level of $\alpha=.01$.

Hypothesis VIC

When the Powell and Betts criteria do not place students at the same instructional level, a greater proportion of

students will be placed at a higher instructional reading level by the Powell criteria than the Betts criteria.

Hypothesis VID

Students will be placed at the same instructional level by the Powell and Betts IRI criteria at least 75 percent of the time.

The number of times the Powell criteria would place students at the same instructional level was calculated by hand and reported in percentages. A Chi Square test was used to test the significance of the differences in proportions of students placed at the higher instructional level by the Powell and Betts criteria, respectively.

CHAPTER IV
RESULTS, DISCUSSION AND
THEORETICAL CONSIDERATIONS

Results and Discussions

The major purpose of this study was to investigate the relationship between standard cloze and MCC tests. Several hypotheses were tested in this regard.

A secondary purpose of the study involved a comparison of readability levels as determined by traditional methods and a new system of readability determination, Rasch calibration (Kidder, 1977). A third aspect of this study involved the IRI and the two types of scoring criteria used with the IRI to determine instructional reading levels.

The results for the testing of each hypothesis will be presented, followed by a discussion of those results. The first hypothesis had the following components:

Hypothesis IA - There will be no differences between the mean scores for standard cloze and MCC passages for the same group of students.

Hypothesis IB - There will be no differences among the mean scores on the three levels of difficulty tested for each group.

Hypothesis IC - There will be no interaction between the form used (standard cloze or MCC) and the level of difficulty (readability level) of the passages.

Mean scores and standard deviations by grade and test are shown in Table 1.

A randomized block factorial design was used to test the four parts of the first hypothesis. Each student was tested on the same six measures. The factors involved in the analysis were test form and difficulty level. Test form consisted of two types of tests, standard cloze and MCC. Passage difficulty consisted of three consecutive grade levels. Separate analyses were conducted for grades two, four, and six. The analysis of variance (ANOVA) tables for these analyses are presented in Tables 2, 3, and 4, respectively.

A significant interaction between test form and passage difficulty was present in grades two, four, and six. Hypothesis IC was rejected. Because the interaction was significant, Simple Main Effects were calculated to determine more accurately where the significant differences were present. Tables 5, 6, and 7 show the Simple Main Effects summary tables for grades two, four, and six, respectively.

The tests for Simple Main Effects at all three grades indicated all effects studied were significant. These findings actually mean that for the level one passage, or the lowest readability level given to each group (first grade level for the second grade group, third grade level for the fourth grade group, and fifth grade level for the sixth grade

Table 1

Means on Standard Cloze and MCC for
Grades Two, Four, and Six

| <u>Grade</u> | <u>Test</u> | <u>Passage Level 1</u> | <u>Passage Level 2</u> | <u>Passage Level 3</u> | <u>N</u> |
|--------------|-------------|----------------------------|----------------------------|----------------------------|----------|
| Grade 2 | SC | 10.6 | 6.7 | 4.0 | 181 |
| Grade 2 | MCC | 15.8 | 17.1 | 13.3 | 181 |
| Grade 4 | SC | 9.5 | 7.9 | 6.4 | 168 |
| Grade 4 | MCC | 20.3 | 19.8 | 17.6 | 168 |
| Grade 6 | SC | 12.4 | 8.7 | 7.3 | 196 |
| Grade 6 | MCC | 21.4 | 21.0 | 18.1 | 196 |

Table 2

Analysis of Variance Abbreviated Summary¹
Table for Grade Two

| Source | df | SS | MS | F |
|------------------------------------------------|-----|----------|----------|----------|
| Test Form | 1 | 18712.32 | 18712.32 | 1049.99* |
| Passage Difficulty Level | 2 | 4019.80 | 2009.90 | 107.30* |
| Test Form by Difficulty Level | 2 | 1297.49 | 648.74 | 32.16* |
| Test Form by Student | 180 | 3207.86 | 17.82 | |
| Difficulty Level by Student | 360 | 6743.10 | 18.73 | |
| Test Form by Difficulty Level by Student | 360 | 7262.51 | 20.17 | |

* p < .05

¹The abbreviated summary table contains only the factors involved in the testing of Hypotheses IA, IB, and IC.

Table 3

Analysis of Variance Abbreviated Summary¹
Table for Grade Four

| Source | df | SS | MS | F |
|------------------------------------------------|-----|----------|----------|----------|
| Test Form | 1 | 32276.84 | 32276.84 | 2118.31* |
| Passage Difficulty Level | 2 | 1472.06 | 736.03 | 104.28* |
| Test Form by Difficulty Level | 2 | 52.71 | 26.36 | 4.29* |
| Test Form by Student | 167 | 2544.59 | 15.24 | |
| Difficulty Level by Student | 334 | 2357.51 | 7.06 | |
| Test Form by Difficulty Level by Student | 334 | 2054.25 | 6.15 | |

* $p < .05$

¹The abbreviated summary table contains only the factors involved in the testing of Hypothesis IA, IB, and IC.

Table 4

Analysis of Variance Abbreviated Summary¹
Table for Grade Six

| Source | df | SS | MS | F |
|------------------------------------------------|-----|----------|----------|----------|
| Test Form | 1 | 33685.16 | 33685.16 | 4060.82* |
| Passage Difficulty Level | 2 | 3545.19 | 1772.59 | 357.71* |
| Test Form by Difficulty Level | 2 | 572.24 | 286.12 | 58.76* |
| Test Form by Student | 195 | 1617.56 | 8.23 | |
| Difficulty Level by Student | 390 | 1932.60 | 4.96 | |
| Test Form by Difficulty Level by Student | 390 | 1902.38 | 4.89 | |

*p < .05

¹The abbreviated summary table contains only the factors involved in the testing of Hypotheses IA, IB, and IC.

Table 5

Simple Main Effects for Grade Two

| <u>Source</u> | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> |
|-----------------|-----------|-----------|-----------|----------|
| Form at Level 1 | 1 | 2519.41 | 2519.41 | 129.93* |
| Form at Level 2 | 1 | 9722.03 | 9722.03 | 501.39* |
| Form at Level 3 | 1 | 7768.86 | 7768.86 | 400.66* |
| Level at SC | 2 | 3944.00 | 1972.00 | 101.39* |
| Level at MCC | 2 | 1373.38 | 686.69 | 35.31* |

*p < .05

Table 6

Simple Main Effects for Grade Four

| <u>Source</u> | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> |
|-----------------|-----------|-----------|-----------|----------|
| Form at Level 3 | 1 | 9793.44 | 9793.44 | 1066.82* |
| Form at Level 4 | 1 | 11916.67 | 11917.67 | 1298.11* |
| Form at Level 5 | 1 | 10620.00 | 10620.00 | 1156.86* |
| Level at SC | 2 | 836.91 | 418.46 | 63.40* |
| Level at MCC | 2 | 687.94 | 343.97 | 52.11* |

p < .05

Table 7

Simple Main Effects for Grade 6

| <u>Source</u> | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> |
|-----------------|-----------|-----------|-----------|----------|
| Form at Level 5 | 1 | 7848.25 | 7848.25 | 1303.70* |
| Form at Level 6 | 1 | 14976.85 | 14976.85 | 2487.85* |
| Form at Level 7 | 1 | 11432.88 | 11432.88 | 1899.15* |
| Level at SC | 2 | 2811.99 | 1406.00 | 285.78* |
| Level at MCC | 2 | 1305.62 | 652.81 | 132.69* |

*p < .05

group), significant differences were observed between student scores on standard cloze tests and MCC tests. The same is also true at the other two grade levels at which the students were tested. The form of the test, standard cloze or MCC, made a significant difference for all three grades at all readability levels examined. Thus, Hypothesis IA was rejected.

For all three samples (second, fourth, and sixth graders) the Simple Main Effects test also indicated significant differences among mean scores on the three passage difficulty levels. Therefore, Hypothesis IB was rejected. Since three passage difficulty levels were tested on each test form, a further test was necessary to determine exactly which level scores were significantly different within each form.

Tukey's test for differences in means was used to make this determination (Hays, 1973). The differences in means were tested for significance at the .05 level. Tukey's test results are presented in Tables 8, 9, and 10.

For the second grade, on the standard cloze test and the MCC test, all of the mean scores were significantly different from each other. In other words, the first level mean scores were significantly different from the second and third level scores, and the second level mean scores were significantly different from the third level mean scores.

For the fourth grade, the standard cloze mean scores were all significantly different on passages of different difficulty. The MCC mean scores on passages of different

Table 8

Tukey's Test for Differences in
Means for Standard Cloze Scores
and MCC Scores for Grade Two

| <u>Standard Cloze</u> | <u>\bar{X}</u> |
|-----------------------|-----------------------------|
| Difficulty Level 1 | 10.56 |
| Difficulty Level 2 | 6.71 |
| Difficulty Level 3 | 3.99 |

| | 1 | 2 | 3 |
|---|---|-------|-------|
| 1 | | 3.85* | 6.57* |
| 2 | | | 2.72* |
| 3 | | | |

$$HSD_{.05} = 3.31 \sqrt{\frac{19.45^1}{181}} = 1.09$$

| <u>Multiple Choice Cloze</u> | <u>\bar{X}</u> |
|------------------------------|-----------------------------|
| Difficulty Level 1 | 15.83 |
| Difficulty Level 2 | 17.07 |
| Difficulty Level 3 | 13.25 |

| | 2 | 1 | 3 |
|---|---|-------|-------|
| 2 | | 1.24* | 3.82* |
| 1 | | | 2.58* |
| 3 | | | |

$$HSD_{.05} = 3.31 \sqrt{\frac{19.45}{181}} = 1.09$$

¹MS ERROR = pooled error terms for difficulty level and test form by difficulty level.

Table 9

Tukey's Test for Differences in Means for Standard Cloze Scores and MCC Scores for Grade Four

| <u>Standard Cloze</u> | <u>\bar{x}</u> |
|-----------------------|-----------------------------|
| Difficulty Level 3 | 9.54 |
| Difficulty Level 4 | 7.87 |
| Difficulty Level 5 | 6.83 |

| | 1 | 2 | 3 |
|---|---|-------|-------|
| 1 | | 1.67* | 3.16* |
| 2 | | | 1.49* |
| 3 | | | |

$$HSD_{.05} = 3.31 \sqrt{\frac{6.6}{168}} = .656$$

| <u>Multiple Choice Cloze</u> | <u>\bar{x}</u> |
|------------------------------|-----------------------------|
| Difficulty Level 3 | 20.33 |
| Difficulty Level 4 | 19.78 |
| Difficulty Level 5 | 17.63 |

| | 1 | 2 | 3 |
|---|---|-----|-------|
| 1 | | .55 | 2.70* |
| 2 | | | 2.15* |
| 3 | | | |

$$HSD_{.05} = 3.31 \sqrt{\frac{6.6}{168}} = .656$$

¹MS ERROR = pooled error terms for difficulty level and test form by difficulty level.

Table 10

Tukey's Test for Differences in
Means for Standard Cloze Scores
and MCC Scores for Grade Six

| <u>Standard Cloze</u> | | <u>\bar{X}</u> |
|-----------------------|--|-----------------------------|
| Difficulty Level 5 | | 12.44 |
| Difficulty Level 6 | | 8.67 |
| Difficulty Level 7 | | 7.27 |

| | 1 | 2 | 3 |
|---|---|-------|-------|
| 1 | | 3.77* | 5.17* |
| 2 | | | 1.40* |
| 3 | | | |

$$\text{HSD}_{.05} = 3.31 \sqrt{\frac{4.92}{196}} = .524$$

| <u>Multiple Choice Cloze</u> | | <u>\bar{X}</u> |
|------------------------------|--|-----------------------------|
| Difficulty Level 5 | | 21.39 |
| Difficulty Level 6 | | 21.03 |
| Difficulty Level 7 | | 18.07 |

| | 1 | 2 | 3 |
|---|---|-----|-------|
| 1 | | .36 | 3.32* |
| 2 | | | 2.96* |
| 3 | | | |

$$\text{HSD}_{.05} = 3.31 \sqrt{\frac{4.92}{196}} = .524$$

¹MS ERROR = pooled error terms for difficulty level and test form by difficulty level.

difficulty were significantly different between levels three and five, and levels four and five.

The results for the sixth grade group for the standard cloze test indicated significant differences between the mean scores for all passage levels.

In the MCC form situation, there were two significant differences between mean scores. The fifth grade level passage mean score was significantly different from the seventh grade level mean scores, and the sixth grade level mean scores were significantly different from the seventh grade level mean scores.

In review, significant differences due to test form were observed at all three grade levels. Significant differences between passages did exist for each test form between most grade levels.

Hypothesis II stated that a correlation of .70 or more exists between standard cloze and MCC. A correlation of .70 or more would be necessary to indicate a practical significance of the relationship of these two variables.

Tables 11, 12, and 13 depict the correlation coefficients for cloze tests at grades two, four, and six, respectively.

Hypothesis II predicted a strong relationship (.70 or more) would exist between standard cloze and MCC. Hypothesis II was not supported at grade four or grade six. For second grade, second level, the correlation was .80. It was expected that the scores on the identical story given to the same students would correlate highly, despite the differences in forms.

Table 11

Second Grade Correlations of Cloze Tests

| | <u>SCT1</u> | <u>SCT2</u> | <u>SCT3</u> | <u>MCCT1</u> | <u>MCCT2</u> | <u>MCCT3</u> | <u>Total SC</u> | <u>Total MCC</u> | <u>Total KC</u> |
|-----------|-------------|-------------|-------------|--------------|--------------|--------------|-----------------|------------------|-----------------|
| SCT1 | 1.00 | .45 | .17 | .45 | .43 | .48 | .59 | .54 | .54 |
| SCT2 | | 1.00 | .81 | .57 | .80 | .49 | .95 | .75 | .62 |
| SCT3 | | | 1.00 | .51 | .74 | .27 | .87 | .62 | .43 |
| MCCT1 | | | | 1.00 | .73 | .30 | .62 | .79 | .59 |
| MCCT2 | | | | | 1.00 | .54 | .82 | .92 | .74 |
| MCCT3 | | | | | | 1.00 | .49 | .78 | .71 |
| Total SC | | | | | | | 1.00 | .78 | .63 |
| Total MCC | | | | | | | | 1.00 | .82 |
| Total KC | | | | | | | | | 1.00 |

Underline indicates practical significance (.70).

Table 12

Fourth Grade Correlations of Cloze Tests

| | <u>SCT1</u> | <u>SCT2</u> | <u>SCT3</u> | <u>MCCT1</u> | <u>MCCT2</u> | <u>MCCT3</u> | Total SC | Total MCC | Total KC |
|--------------|-------------|-------------|-------------|--------------|--------------|--------------|-------------|--------------|-------------|
| SCT1 | 1.0 | <u>.75</u> | .62 | .62 | .60 | .61 | <u>.87</u> | .68 | <u>.71</u> |
| SCT2 | | 1.0 | <u>.79</u> | .64 | .64 | .64 | <u>.94</u> | <u>.71</u> | .67 |
| SCT3 | | | 1.0 | .51 | .53 | .56 | <u>.90</u> | .59 | .52 |
| MCCT1 | | | | 1.0 | <u>.77</u> | .64 | .65 | <u>.87</u> | <u>.74</u> |
| MCCT2 | | | | | 1.0 | <u>.76</u> | .65 | <u>.93</u> | <u>.72</u> |
| MCCT3 | | | | | | 1.0 | .67 | <u>.91</u> | <u>.71</u> |
| Total SC | | | | | | | 1.0 | <u>.73</u> | <u>.70</u> |
| Total MCC | | | | | | | | 1.0 | <u>.80</u> |
| Total KC | | | | | | | | | 1.0 |

Underline indicates practical significance (.70).

Table 13

Sixth Grade Correlations of Cloze Test Scores

| | <u>SCT1</u> | <u>SCT2</u> | <u>SCT3</u> | <u>MCCT1</u> | <u>MCCT2</u> | <u>MCCT3</u> | <u>Total SC</u> | <u>Total MCC</u> | <u>Total KC</u> |
|--------------|-------------|-------------|-------------|--------------|--------------|--------------|---------------------|----------------------|---------------------|
| SCT1 | 1.0 | .65 | .62 | .46 | .56 | .56 | <u>.89</u> | .63 | .62 |
| SCT2 | | 1.0 | .59 | .39 | .47 | .42 | <u>.86</u> | .51 | .51 |
| SCT3 | | | 1.0 | .35 | .53 | .56 | <u>.84</u> | .58 | .51 |
| MCCT1 | | | | 1.0 | .53 | .44 | .47 | <u>.76</u> | .60 |
| MCCT2 | | | | | 1.0 | .67 | .60 | <u>.90</u> | <u>.73</u> |
| MCCT3 | | | | | | 1.0 | .60 | <u>.84</u> | .63 |
| Total SC | | | | | | | 1.0 | .67 | .64 |
| Total MCC | | | | | | | | 1.0 | <u>.79</u> |
| Total KC | | | | | | | | | 1.0 |

Underline indicates practical significance (.70).

It should be noted, however, correlations between total scores over three passages were .78, .73, and .67 for grades two, four, and six, respectively.

The fact that the correlations for the same stories were so low, in some instances indicating a shared variance of less than 8 percent, suggests that standard cloze and MCC may not measure the same thing. A shared variance of less than 8 percent indicates that less than 8 percent of the variance in scores of MCC can be explained by variance in standard cloze scores.

To further investigate the relationship between MCC test scores and standard cloze test scores the correlational information was examined using the multitrait-multimethod matrix suggested by Campbell and Fisk (1959).

The MCC and standard cloze score correlations for the same passage level were considered to be correlations between measures of the same trait using different methods. These correlations are called the convergent validity coefficients. Campbell and Fisk reason that such correlations should be greater than the correlations between measures of different traits with different methods (i.e., MCC and standard cloze score correlations for different passage levels).

The second grade correlations for the diagonal (same passage-different method) ranged from .46 to .56. If the two methods (MCC and standard cloze) measured the same construct, these correlations should exceed the correlations located within the dotted line triangles. Similarly, diagonal

correlations should at least equal the values contained within the solid triangles (correlations between different passages using same method). These ranged from .44 to .69. The different method-different passage correlations ranged from .39 to .56.

The ranges of the correlations in Table 13 were similar enough to question whether standard cloze and MCC measure the same or independent constructs. The second grade correlations were so close in range that the question of whether standard cloze and MCC both measure reading comprehension cannot be proven or disproven.

The fourth grade test score correlations for different method-same passage ranged from .56 to .64. The same method-different passage correlations ranged from .62 to .79. The different method-different passage correlations ranged from .51 to .64. At this grade level also the range of correlations were too similar to determine independence or strong relationship of measured constructs.

The sixth grade test score correlations for different method-same passage ranged from .46 to .56. The same method-different passage correlations ranged from .44 to .67. The different method-different passage correlations ranged from .35 to .56. In this case, also, the ranges of the correlations were too close to conclude whether standard cloze and MCC were measuring like or different constructs. The similarities of the range of correlations also preclude any strong statements reflecting that standard cloze and MCC do or do not both

measure reading comprehension, or the same type of reading comprehension. Additional study in this area might help to clarify this situation.

Guthrie (1974) suggested the use of Maze Technique (a MCC form developed in the same way as Homan MCC) to determine reading comprehension and if a book is at a suitable instructional level for a student. Standard cloze has long been regarded as a measure of reading comprehension (Bormuth, 1969; Horton, 1975; Rankin, 1959). Some question has been raised, as evidenced by the current study, concerning whether or not MCC and standard cloze tests measure the same thing. Perhaps Guthrie's contention that Maze or MCC measure reading comprehension should be reevaluated and investigated further.

Informal Reading Inventory scores correlated more highly with standard cloze scores than MCC scores (Table 13) at the sixth grade level, again indicating that MCC or Maze may not be measuring reading comprehension.

Guthrie (1974) suggested that the amount of items correct is the student's percentage of amount of comprehension on any given Maze passage. In light of the present study, this seems unlikely. Additional study is called for to determine what MCC is measuring. If it is measuring reading comprehension, is it a different type of comprehension than that measured by standard cloze tests and the IRI?

The fourth grade correlations between standard cloze and MCC scores were the highest of the three grade levels. However, even these scores fell short of the .70 value of

practical significance. Most of the correlations at all three grades indicated approximately 28 percent to 30 percent shared variance.

The value of .70 was chosen because it would indicate 49 percent shared variance; that is, almost 50 percent of the variance in MCC scores could be explained by standard cloze scores. If they had less variance in common, it would not seem prudent to set up a new criteria for determining instructional level with MCC from standard cloze scores.

The total score correlations were much higher than those of the individual test scores. The total scores were obtained by using combined total scores across the three levels tested at each grade. For example, at the second grade level the total standard cloze score was the combined score on passages one, two, and three.

These total correlations were about .70 in both second and fourth grades. The correlation neared practical significance in sixth grade with a .67 correlation between standard cloze and MCC scores. This suggests some value in giving three passages to each student. The time involved in the testing procedure would be tripled. By extending the testing time, much of the value of standard cloze and MCC would be lost for the classroom teacher.

Null Hypothesis III was in three parts. In the first two parts it stated that there would be no relationship between standard cloze scores and IRI scores or between IRI and MCC scores.

Table 14 contains the IRI, standard cloze, and MCC correlations for the second, fourth, and sixth grades. The Powell differentiated criteria (Powell, 1978) were used to score the IRI's.

Hypothesis IIIA stated there would be no relationship between standard cloze scores and IRI scores. Hypothesis IIIB stated the same thing for MCC scores and IRI scores. However, in many instances, that was not the case.

In the second grade data the IRI correlated significantly with the standard cloze level one passage and the MCC third level passage. The other correlations were not significant at the $\alpha=.01$ level.

None of the fourth grade tests, standard cloze or MCC, correlated significantly with the IRI scores. In the sixth grade the fifth level MCC did not correlate significantly with the IRI scores. The seventh level MCC scores were very close to a significant correlation. All other sixth grade tests did correlate significantly with the IRI scores ($\alpha=.01$).

Hypotheses IIIA and IIIB were rejected for grades two and six, but these null hypotheses were not rejected for grade four.

The third part of Hypothesis III stated that correlations between MCC scores and IRI instructional level scores would be the same as correlations between standard cloze and IRI instructional level scores. The significance of the differences in correlations was determined by using the formula for comparison of correlation coefficients from the same sample (Guilford & Fruchter, 1973; Hotelling, 1940).

Table 14

Correlations of Informal Reading Inventory
Instructional Level Scores with Standard
Cloze Scores and Multiple Choice Cloze
Scores for Grades Two, Four, and Six

| <u>Grade Two</u> | <u>N</u> | <u>IRI</u> | <u>SCT1</u> | <u>SCT2</u> | <u>SCT3</u> | <u>MCCT1</u> | <u>MCCT2</u> | <u>MCCT3</u> |
|-------------------|----------|------------|-------------|-------------|-------------|--------------|--------------|--------------|
| IRI | 33 | 1.0 | <u>.81</u> | .32 | .07 | .16 | .12 | <u>.75</u> |
| <u>Grade Four</u> | <u>N</u> | <u>IRI</u> | <u>SCT3</u> | <u>SCT4</u> | <u>SCT5</u> | <u>MCCT3</u> | <u>MCCT4</u> | <u>MCCT5</u> |
| IRI | 30 | 1.0 | .25 | .33 | .32 | .27 | .21 | .17 |
| <u>Grade Six</u> | <u>N</u> | <u>IRI</u> | <u>SCT5</u> | <u>SCT6</u> | <u>SCT7</u> | <u>MCCT5</u> | <u>MCCT6</u> | <u>MCCT7</u> |
| IRI | 31 | 1.0 | <u>.51</u> | <u>.52</u> | <u>.43</u> | -.08 | <u>.55</u> | .38 |

Underline indicates $p < .01$.

$$t = (r_{12} - r_{13}) \sqrt{\frac{(N-3)(1+r_{23}^2)}{2(1-r_{23}^2 - r_{12}^2 - r_{13}^2 + 2r_{23}r_{13})}}$$

where,

r_{12} is the correlation between IRI and MCC.

r_{13} is the correlation between IRI and standard cloze.

r_{23} is the correlation between MCC and standard cloze.

d.f. is $N-3$.

The results of this test for significance were mixed. In the second grade standard cloze scores correlations with the IRI were significantly different from MCC correlations with the IRI at all three passages ($\alpha=.01$). However, in passage levels one and two, standard cloze correlations were higher, while in passage level three, the MCC correlation was higher. This was possibly due to the extremely low scores many second graders received on the standard cloze third grade level passage. It was very frustrating for many of them.

Hypothesis IIIC was not accepted for all three passages of the second grade test. It should be noted, however, that standard cloze correlations were higher at levels one and two.

The fourth grade passages showed no significant differences ($\alpha=.01$) at any passage level. Therefore, Hypothesis IIIC was accepted for the fourth grade level passages. There was no significant difference between standard cloze correlations to IRI scores and MCC correlations to the IRI.

The sixth grade results were mixed. On the fifth grade level passage the standard cloze correlation with IRI was significantly greater than the MCC correlation to the IRI. However, on the other two passages (level six and seven) the correlations were not significantly different ($\alpha=.01$). Hypothesis IIIC was not accepted for level five of the sixth grade passages. It was accepted for levels six and seven where no significant differences were indicated between correlations of MCC and IRI scores and standard cloze and IRI scores.

Hypothesis IV states that MCC passage reliabilities will be the same as standard cloze passage reliabilities at the same grade level. Tables 15, 16, and 17 list the passage reliabilities, the range of item difficulties, biserial correlations, and standard error of measurement for the second, fourth, and sixth grade tests, respectively.

These results indicated Hypothesis IV was not accepted. At all levels for all three grades the MCC test reliabilities were higher than the standard cloze test reliabilities at the same passage level. All passages at all levels showed high reliabilities. However, the sixth grade passages, in some instances, had lower reliabilities than the second and fourth grade passages.

The reliabilities were a measure of internal consistency. The high reliabilities indicated that a large number of items on a test were measuring the same thing.

Table 15

Passage Reliabilities, Standard Errors of
Measurement, Item Difficulties, and
Biserial Correlations for all Cloze
Tests for the Second Grade

| | <u>Reliability</u> | <u>SE</u> | <u>Range of Item Difficulty</u> | <u>Range of Biserial Correlations</u> |
|--------|--------------------|-----------|-----------------------------------------|-----------------------------------------------|
| SCT1 | .87 | 1.61 | .19 - .83 | .46 - 1.0 ^a |
| SCT2 | .89 | 1.56 | .00 - .65 | .00 - .96 |
| SCT3 | .89 | 1.30 | .00 - .38 | .00 - .99 |
| HMCCT1 | .93 | 1.04 | .71 - .92 | .32 - 1.0 ^a |
| HMCCT2 | .94 | 1.61 | .57 - .87 | .65 - 1.0 ^a |
| HMCCT3 | .91 | 1.87 | .28 - .69 | .50 - .98 |
| KCT1 | .90 | 1.50 | .54 - .86 | .55 - 1.0 ^a |
| KCT2 | .89 | 1.25 | .57 - .81 | .61 - 1.0 ^a |
| KCT3 | .80 | 1.21 | .42 - .75 | .56 - .92 |
| TKC | .95 | 2.43 | .42 - .86 | .47 - .98 |

^aValue larger than 1.0 possibly due to a violation of the assumption of a normal distribution.

Table 16

Passage Reliabilities, Standard Errors of Measurement, Item Difficulties, and Biserial Correlations for all Cloze Tests for the Fourth Grade

| | <u>Reliability</u> | <u>SE</u> | <u>Range of Item Difficulty</u> | <u>Range of Biserial Correlations</u> |
|--------|--------------------|-----------|---------------------------------|---------------------------------------|
| SCT3 | .83 | 1.69 | .00 - .84 | .00 - .95 |
| SCT4 | .85 | 1.65 | .00 - .73 | .00 - .95 |
| SCT5 | .89 | 1.58 | .00 - .66 | .00 - .97 |
| HMCCT3 | .91 | 1.22 | .70 - .95 | .63 - 1.0 ^a |
| HMCCT4 | .92 | 1.32 | .65 - .95 | .68 - 1.0 ^a |
| HMCCT5 | .92 | 1.58 | .49 - .87 | .38 - 1.0 ^a |
| KCT3 | .79 | .80 | .73 - .95 | .80 - 1.0 ^a |
| KCT4 | .78 | .99 | .63 - .91 | .59 - 1.0 ^a |
| KCT5 | .79 | .90 | .58 - .94 | .66 - 1.0 ^a |
| TKC | .91 | 1.63 | .58 - .95 | .45 - 1.0 ^a |

^aValue larger than 1.0 due possibly to a violation of the assumption of a normal distribution.

Table 17

Passage Reliabilities, Standard Errors of
Measurement, Item Difficulties, and
Biserial Correlations for all Cloze
Tests for the Sixth Grade

| | <u>Reliability</u> | <u>SE</u> | <u>Range of Item Difficulty</u> | <u>Range of Biserial Correlations</u> |
|--------|--------------------|-----------|-----------------------------------------|-----------------------------------------------|
| SCT5 | .75 | 1.80 | .00 - .93 | .00 - .91 |
| SCT6 | .70 | 1.80 | .01 - .94 | .06 - .84 |
| SCT7 | .69 | 1.65 | .00 - .82 | .00 - .77 |
| HMCCT5 | .88 | 1.00 | .68 - .99 | .65 - 1.0 ^a |
| HMCCT6 | .92 | 1.06 | .83 - .96 | .62 - 1.0 ^a |
| HMCCT7 | .84 | 1.28 | .40 - .96 | .30 - 1.0 ^a |
| KCT5 | .67 | .54 | .86 - .99 | .86 - 1.0 ^a |
| KCT6 | .82 | .49 | .87 - .99 | .96 - 1.0 ^a |
| KCT7 | .65 | .98 | .21 - .98 | .35 - 1.0 ^a |
| TKC | .85 | 1.28 | .21 - .99 | .21 - 1.0 ^a |

^aValue larger than 1.0 due possibly to a violation of the assumption of a normal distribution.

In all three grades, for all passages, the Homan MCC passages had higher reliabilities than the standard cloze passages. Hypothesis IV was not supported. Homan MCC passage reliabilities were higher than standard cloze passage reliabilities.

The range of item difficulties in many instances (Tables 15, 16, and 17) was very broad. In some cases, the MCC item difficulties were very high. The sixth grade MCC item difficulties for passages five and six ranged from .86 to .99. This indicates that all the items were too easy and that almost all students were able to get almost all items correct. When many students get very high scores it is very difficult to differentiate between students' abilities.

The biserial correlations between the individual item scores and the total test scores were varied. They tended to cover the total range of possible correlations.

Hypothesis V in the null form stated that there was no relationship between total Homan MCC scores and total Kidder MCC scores. Pearson Product-Moment correlations were used to determine the strength of the relationship. Table 18 presents the results of these correlations for second, fourth, and sixth grades.

Hypothesis V was rejected for all grade levels. A significant relationship ($\alpha=.01$) did exist between total Homan MCC scores and total Kidder MCC scores.

Total scores were used in this correlation due to the shortness of some of the Kidder passages. In many instances

Table 18

Correlations of Homan Multiple Choice Cloze
Total Scores and Kidder Multiple Choice
Total Scores for the Second, Fourth,
and Sixth Grades

| | Grade 2 Total <u>HMCC</u> | Grade 2 Total <u>KMCC</u> | Grade 4 Total <u>HMCC</u> | Grade 4 Total <u>KMCC</u> | Grade 6 Total <u>HMCC</u> | Grade 6 Total <u>KMCC</u> |
|--------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Grade 2 Total HMCC | 1.0 | <u>.82</u> | | | | |
| Grade 2 Total KMCC | - | 1.0 | | | | |
| Grade 4 Total HMCC | - | - | 1.0 | <u>.80</u> | | |
| Grade 4 Total KMCC | - | - | - | 1.0 | | |
| Grade 6 Total HMCC | - | - | - | - | 1.0 | <u>.79</u> |
| Grade 6 Total KMCC | - | - | - | - | - | 1.0 |

Underline indicates $\alpha = .01$.

a test consisted of only ten items. Also, the reliabilities for the Kidder cloze passages individually (Tables 15, 16, and 17) were lower than the total Kidder cloze passage reliabilities. However, in no instance did an individual test have a reliability lower than .65.

The Kidder cloze passages had been Rasch calibrated to determine readability level (Kidder, 1977). High correlations between Kidder MCC and Homan MCC passages indicate a strong relationship between readability levels as determined by traditional formulas (Dale-Chall, 1948; Harris-Jacobson in Harris & Sipay, 1975; Spache, 1974) and the new method of Rasch calibration.

The significant high correlations between Homan MCC and Kidder MCC also suggest that no learning took place for the students from the first testing of standard cloze passages to the second testing with Homan MCC. If the students had experienced information gain from the standard cloze tests, it would have been evidenced by low correlations between Kidder MCC and Homan MCC, since half of the sample were given the Kidder MCC passages at the first sitting, and the other half were administered the Kidder MCC passages at the second sitting.

Hypothesis VIA stated that there would be no relationship between IRI's instructional level scores as scored by the Powell and Betts criteria ($\alpha=.01$). Hypothesis VIB stated there would be no difference between instructional level means as scored by the Powell criteria and Betts criteria ($\alpha=.01$).

Hypothesis VIC stated that when the Powell and Betts criteria did not place students at the same instructional level, a greater proportion of students would be placed at a higher instructional reading level by the Powell criteria than the Betts criteria ($\alpha=.01$). Hypothesis VID stated that students would be placed at the same instructional level by the Powell and Betts IRI criteria at least 75 percent of the time.

Table 19 shows the correlations between IRI's scored by the Powell criteria and IRI's scored by the Betts criteria at the second, fourth, and sixth grade levels.

Hypothesis VIA was rejected at all three grade levels. There was a significant ($\alpha=.01$) relationship between IRI's scored by the Powell criteria and IRI's scored by the Betts criteria.

Hypothesis VIB involved the differences between instructional level mean scores when an IRI was scored by the Powell criteria and Betts criteria. Table 20 depicts the means and standard deviation of the Powell criteria instructional level means and the Betts criteria instructional level means.

At all grade levels the Powell criteria instructional level means were higher than the Betts criteria instructional level means. The differences between the means, when tested for significance with a t-test for means, were all significant at the .01 and the .001 levels. In terms of placement of students, this means an important difference exists depending on whether the Powell or Betts criteria are used to score the IRI.

Table 19

Correlations of the Informal Reading Inventory's
Instructional Level Scores as Scored by the
Powell Criteria and Betts Criteria for
the Second, Fourth, and Sixth Grades

| | Grade 2 Powell IRI | Grade 2 Betts IRI | Grade 4 Powell IRI | Grade 4 Betts IRI | Grade 6 Powell IRI | Grade 6 Betts IRI |
|--------------------------|--------------------------|-------------------------|--------------------------|-------------------------|--------------------------|-------------------------|
| Grade 2 Powell IRI | 1.0 | <u>.80</u> | | | | |
| Grade 2 Betts IRI | - | 1.0 | | | | |
| Grade 4 Powell IRI | - | - | 1.0 | <u>.80</u> | | |
| Grade 4 Betts IRI | - | - | | 1.0 | | |
| Grade 6 Powell IRI | - | - | - | - | 1.0 | <u>.85</u> |
| Grade 6 Betts IRI | - | - | - | - | - | 1.0 |

Underline indicates $\alpha = .001$.

Table 20

Means, Standard Deviations, and Standard Errors of Instructional Level Scores as Determined by the Powell Criteria and Betts Criteria for the Second, Fourth, and Sixth Grades

| | <u>N</u> | <u>\bar{X}</u> | <u>SD</u> | <u>SE</u> |
|-----------------------|----------|-----------------------------|-----------|-----------|
| Grade 2 Powell IRI | 32 | 3.0313 | 1.656 | 0.293 |
| Grade 2 Betts IRI | 32 | 2.1875 | 1.712 | 0.303 |
| Grade 4 Powell IRI | 30 | 4.533 | 1.224 | 0.224 |
| Grade 4 Betts IRI | 30 | 3.867 | 1.479 | 0.270 |
| Grade 6 Powell IRI | 31 | 5.419 | 1.177 | 0.211 |
| Grade 6 Betts IRI | 31 | 4.903 | 1.535 | 0.276 |

Hypothesis VIB was rejected at all levels. A significant difference in instructional level means did exist between IRI scores obtained by the Powell criteria and those obtained by the Betts criteria.

While investigating these differences, additional information concerning instructional level means became apparent. A comparison of instructional level means using the Powell and Betts criteria was made at grade levels one to seven.

Table 21 shows the instructional level standard cloze percentages as determined from IRI's scored by the Powell and Betts criteria.

Fewer students were at each instructional level using the Betts criteria. More students tended not to fit in at any one level. Due to the nature of the cloze passages used, second grade students would need to have scored at the first, second, or third grade instructional level on the IRI to have their scores included in the instructional level means. Fourth graders had to be instructional at the third, fourth, or fifth grade level, and sixth graders had to be instructional at the fifth, sixth, or seventh grade levels to be included.

The range of standard cloze instructional level scores using the Powell criteria was 13 percent to 45 percent. The 13 percent for second grade was exceptionally low. This percentage was probably distorted because five of the thirteen students received scores of zero on the standard cloze tests. Many of them did not even attempt to complete any of the

Table 21

Instructional Level Cloze Score Percentages
as Determined by the Powell Criteria
and Betts Criteria

| <u>Grade</u> | <u>N</u> | <u>Cloze % Using Betts Criteria</u> | <u>N</u> | <u>Cloze % Using Powell Criteria</u> |
|--------------|----------|-----------------------------------------|----------|------------------------------------------|
| 1 | 5 | 57% | 0 | - |
| 2 | 7 | 36% | 13 | 13% |
| 3 | 11 | 37% | 7 | 37% |
| 4 | 11 | 39% | 9 | 36% |
| 5 | 10 | 48% | 20 | 45% |
| 6 | 12 | 38% | 11 | 38% |
| 7 | <u>3</u> | 41% | <u>5</u> | 41% |
| Total | 59 | | 65 | |

items. Ignoring that percentage, the range of instructional level scores using the Powell criteria was 36 to 45 percent. The range using the Betts criteria was 36 to 57 percent. Using either IRI scoring criteria base, it indicates that the existing criteria for standard cloze instructional level scores may be too high. The original Bormuth (1967) criteria of 38 to 50 percent comes closest to matching the percentages derived from the present study. The other criteria, Bormuth (1968) of 44 to 57 percent, Rankin and Culhane (1969) from 41 to 61 percent, and Alexander (1968) from 47 to 60 percent, all seem too high. Perhaps criteria differentiated by grade level are needed.

The closeness in range of Powell and Betts cloze instructional scores did not reflect the very real differences present in placing students at their instructional reading level based on the Powell or Betts criteria.

Hypothesis VIC involved the placement of students at instructional level by the Powell and Betts criteria. The hypothesis stated that when the Powell and Betts criteria do not place students at the same instructional level, a significantly greater proportion of students will be placed at a higher instructional reading level by the Powell criteria than the Betts criteria ($\alpha=.01$).

A Chi Square test was done to test this hypothesis. It was expected that when the Powell and Betts criteria placed students at different levels the Powell criteria would place students at a higher instructional level 50 percent of the time.

Hypothesis VIC was accepted. The Chi Square test results indicated that the number of times the Powell criteria placed a student at a higher instructional level than the Betts criteria was significant. In fact, in every case where the Powell and Betts criteria did not place students at the same instructional level, the Powell criteria placed students at a higher level.

Hypothesis VID stated that the Powell and Betts criteria would place students at the same instructional level at least 75 percent of the time. Using percentages, it was determined that the Powell and Betts criteria placed students at the same instructional level only 51 percent of the time. Forty-nine percent of the time the Powell and Betts criteria placed students at different instructional levels.

Of that 49 percent, the Powell criteria always placed students at a higher instructional level. Fifty-five percent of the time the Powell criteria placed the students one grade level above the Betts criteria placement. Thirty-eight percent of the time the Powell criteria placed the student two grade levels above the Betts criteria placement. However, in five cases (28 percent) the two levels' differences involved a jump from frustration level according to the Betts criteria, to the second grade instructional level according to the Powell criteria. In the remaining 7 percent (three cases) the Powell criteria placed the student three grade levels above the Betts criteria.

The highest instructional level score attained by a student was used for analysis in all parts of Hypothesis VI. Hypothesis VID is rejected. The Powell and Betts criteria did not place students at the same instructional level 75 percent of the time.

The important and statistically significant differences in placement based on scoring criteria indicate that classroom teachers should choose the IRI scoring criteria they will use carefully based on a conviction of accuracy of placement.

Theoretical Considerations

One of the original intents of this study was to set up a criteria for Homan MCC to be used by the classroom teacher for determining instructional reading level. This was to be accomplished by using a regression equation based on Homan MCC scores regressed on standard cloze scores for the purpose of setting up an equation that would predict Homan MCC scores. If the predicted Homan MCC scores correlated highly with the actual Homan MCC scores, criteria for Homan MCC instructional level scores could be set up based on how the same students performed on standard cloze tests.

It was assumed that the correlations between standard cloze and MCC would be high, especially since the same students were given the same passages, only in different form.

However, the correlations between standard cloze and MCC scores were much lower than anticipated. The accuracy and practicality of using standard cloze to predict Homan MCC

cloze scores became an issue. Since only the second grade, second level correlation scores reached the desired level of practical significance, it was impractical to set up the new criteria for Homan MCC. The criteria would not have been accurate for classroom use.

The fourth grade correlations represented the highest overall correlations between standard cloze and MCC. Also, the fourth grade distributions were all normal. For this reason, the fourth grade scores were used to determine how accurate or practical a prediction of scores could be made from standard cloze to Homan MCC.

The fourth grade sample was randomly divided into halves. One half was used to derive the regression weights; the other, to cross validate the regression weights. This was done with three separate regressions, one for each passage level. The regression equations are listed in Table 22.

The predicted Homan MCC scores based on the regression equation were then correlated with the actual Homan MCC scores of the second half of the sample. The correlations ranged from .61 to .69. Using the highest correlational .69, only 47.6 percent shared variance existed between standard cloze scores and Homan MCC scores.

The purpose of the prediction equation was to set up the criteria for determining instructional reading level. When dealing with the placement of students in the correct instructional level and thereby correct reader, having more than 50

Table 22

Constant and b Values Used
in the Prediction Equation

| <u>Grade Level</u> | <u>b</u> | <u>Constant</u> |
|--------------------|----------|-----------------|
| 3 | 0.4837 | 16.1227 |
| 4 | 0.6950 | 14.4465 |
| 5 | 0.6091 | 13.9831 |

percent of the variance of MCC not explained by standard cloze leaves too much room for placement error.

These results reinforce that standard cloze scores should not be used to develop a criteria for using Homan MCC scores in the classroom to determine instructional reading level.

CHAPTER V
SUMMARY AND CONCLUSIONS

Standard cloze and the IRI are both used to determine instructional reading level. This study was designed to explore the relationship of a new form of cloze test, MCC, to standard cloze and the IRI. The intent has been to obtain new information that would assist the classroom teacher in determining the instructional reading level of all students as easily and accurately as possible.

Six main objectives are included in the investigation.

1. Comparison of standard cloze and MCC tests for the same passages, with the same students.

2. Examination of the strength of the relationship between standard cloze and MCC passages for the purpose of determining if standard cloze scores can be used to predict MCC scores.

3. Investigation of the relationship among standard cloze scores, MCC scores, and the IRI.

4. Determining the reliabilities of the standard cloze, MCC, and Kidder MCC passages.

5. Examination of the relationship between two forms of MCC tests, Homan MCC and Kidder MCC.

6. Investigation of the relationship between IRI's instructional level scores as scored by the Powell criteria and Betts criteria.

Three schools, representing high, medium, and low socio-economic groups were involved in the study. All second, fourth, and sixth graders at all three schools participated in the study. A total of 548 students were included in the final analysis of data.

Each student was administered three standard cloze tests. One test was at the student's grade level, one was a grade level above the student's grade level, and the third was a grade level below the student's grade level.

Multiple choice cloze tests were made from the same passages used for the standard cloze tests. On a second testing occasion, each student was administered three MCC tests covering the same grade levels as the standard cloze tests.

Additional Kidder MCC tests were administered. These tests were also at three different grade levels. One-half of the students were administered the Kidder MCC test with the standard cloze testing, and the other half were administered the Kidder MCC test with the Homan MCC test. Individual IRI's were administered to at least thirty students at each grade level.

Six hypotheses were tested in the study. Some hypotheses had several sub-hypotheses connected to them.

1. Hypothesis IA stated that there would be no differences between the mean scores for standard cloze and MCC

passages for the same group of students. Mean scores, however, were not equal, and the differences were shown to be statistically significant. The form of cloze test used did affect the score of the students. On the average, students scored lower on standard cloze tests than on MCC tests.

Hypothesis IB stated that there would be no differences among the mean scores on the three levels of difficulty tested for each group. For the second grade all passage scores were significantly different from each other. At the fourth grade level all standard cloze mean scores were significantly different from each other. However, for MCC scores the fifth grade passage mean scores were significantly different from both the third and fourth grade level mean scores. The third and fourth level mean scores were not significantly different from each other. At the sixth grade level all standard cloze mean scores were significantly different from each other. The sixth grade MCC scores indicated the seventh level passage scores were significantly different from the fifth and sixth level passage scores. As would be expected, student scores decreased as the readability levels of the passages increased. The MCC passages did not also show this difference in scores, possibly because of the very high scores many students received on all three levels of MCC passages.

Hypothesis IC stated there would be no interaction between the form used and the level of passage difficulty. A

significant interaction was evidenced by the data, and did, in fact, exist.

2. It was hypothesized that a correlation of .70 or more exists between standard cloze and MCC. Surprisingly, this was not the case. The correlations between standard cloze and MCC on the same passages were much lower than anticipated. Thus, it seems reasonable to question whether the two test forms actually measure the same type of reading comprehension.

3. It was hypothesized in IIIA that there would be no relationship between standard cloze scores and IRI scores. Hypothesis IIIB stated there would be no relationship between MCC and IRI scores. These correlations were very low in some instances and only significant on occasion. The sixth grade correlations were most often significant. This raises questions as to whether the IRI and standard cloze are both measuring the same thing in attempting to determine instructional level.

Hypothesis IIIC stated the correlations between MCC scores and IRI instructional level scores will be the same as the correlations between standard cloze scores and IRI instructional level scores. The results of analysis were mixed.

For the second grade sample significant differences were apparent between correlations of IRI scores and MCC scores and standard cloze and IRI scores. At two passage levels the standard cloze and IRI score correlations were higher.

The fourth grade passage scores showed no significant differences. However, the sixth grade passage scores evidenced mixed results. On the fifth grade level passage the standard cloze correlation with IRI scores was significantly higher than the MCC and IRI score correlation. The sixth and seventh grade level passage scores indicated no significant differences between correlations of standard cloze and IRI scores and MCC and IRI scores. Since the degree of correlation between standard cloze and MCC to the IRI was not significant in many instances, it would be inappropriate to try to draw an overall conclusion about which test form is most closely correlated with the IRI.

4. It was hypothesized that MCC passage reliabilities would be the same as standard cloze passage reliabilities at the same passage level. In fact, MCC and standard cloze passage reliabilities were not the same. MCC reliabilities were higher for all grade levels on all passages.

5. Hypothesis V stated that there would be no relationship between total Homan MCC scores and total Kidder MCC scores. The correlations were significant at all levels. This indicated also that traditional readability formulas and Rasch calibrated readability measure the same thing.

6. Hypothesis VI had four parts. Hypothesis VI stated there would be no relationship between IRI's scored by the Powell and Betts criteria. A significant relationship did exist between the results of the two scoring criteria.

However, Hypotheses VIB, VIC, and VID dealt with differences between the two scoring criteria and found them to be significant and of vital importance concerning placement. There were significant differences at all three grade levels between the mean scores obtained by the two different scoring criteria.

The two sets of criteria place the student at the same level only 51 percent of the time. Furthermore, the analysis indicated that when the two sets of criteria did not place the students at the same level, the Powell criteria placed the students at a higher level 100 percent of the time.

This empirical examination of the relationship among standard cloze, and IRI, and MCC, yielded several useful findings. First, there is some evidence that MCC does not measure the same type of reading comprehension as measured by standard cloze or the IRI. The information on this issue is inconclusive and further study is indicated before more specific conclusions can be reached. Second, the relationship between standard cloze and the IRI also appears tenuous. Third, the scoring criteria used for the IRI can make very significant differences in terms of accurate student placement. While these findings are not definitive for the classroom teacher, they add to the much needed specific information concerning testing and placement for instructional reading level.

APPENDIX A
DIRECTIONS FOR CLOZE TESTINGS

DIRECTIONS FOR STANDARD CLOZE TESTING

Boys and girls, this exercise is trying to find a new way to be able to tell if you understand what you read. Every fifth word has been left out of the story. You must think of a word that makes sense in the story and write it on the line. All lines are the same length. A big word might be missing or a little word like "a," and "the." Only one word can go in each space.

Second Grade Example

The dog ran to _____ door. He wanted to _____
in.

Fourth and Sixth Grade Example

The swing went back _____ forth.

If you leave some blanks during the first reading (and you probably will) go back and try them again.

If you can't spell a word, hold up your hand and I will help you. I can't tell you any words, but I can help you spell words.

When you finish one story start working on the next one. Continue working until all stories are finished. Raise your hand when you're finished and your test will be collected.

Are there any questions?

You may begin.

DIRECTIONS FOR ALL
MULTIPLE CHOICE CLOZE TESTINGS

Here are some stories for you to read. This time there are no blank spaces. But, for every fifth word you'll see three words in one space. Circle the one word that makes the most sense in the story. There is only one correct answer.

Are there any questions?

You may begin.

APPENDIX B
STANDARD CLOZE PASSAGES

STANDARD CLOZE

Test 1

Look at the cat. The cat is very fat. The children liked the _____ cat. The fat cat _____ away. The children looked _____ the fat cat. The _____ cat looked for the _____. The fat cat was _____. Did you see a _____ cat?

The fat cat _____ back. He looked for _____ children. He looked this _____. He looked that way. _____ children looked too. They _____ for the fat cat. _____ fat cat said, "Help!" _____ children ran. They ran _____ the fat cat. He _____ not lost now. The _____ cat was home. He _____ happy.

Test 2

Once upon a time a king, a queen, and their son Todd lived on the top of a huge mountain. They lived in a _____ castle.

There was something _____ wanted very much. He _____ a puppy. He knew _____ would be fun to _____ with a dog.

One _____ Todd asked the king _____ queen for a dog. _____ king thought about it _____ a long time, but _____ wouldn't give Todd a _____ answer. This made Todd _____.

Several days passed. Todd _____ thought about a puppy. _____ next day a large _____ drove up to the _____ entrance. The driver walked _____ to the door.

"I _____ something for Todd," he _____. The man took a _____ out of the car. _____! There were three happy _____ puppies in the box. _____ Todd was really happy.

Test 3

Arnold was not an ordinary cat. His magic powers made _____ very special. It was _____ to spot Arnold in _____ crowd of cats. He _____ the one who winked _____ eye and stuck his _____ in his mouth. If _____ winked with his left _____ and had his tail _____ his mouth it meant _____ unusual was going to _____, however, when Arnold winked _____ his right eye and _____ his tail in his _____ it meant something else. _____ meant something strange, possibly _____ frightening, was about to _____.

One day as I _____ walking down the street _____ saw Arnold step out _____ between two garbage cans. _____ decided to follow Arnold _____ see if he was _____ on using his powers.

Test 4

It was a cold, damp, and rainy day. Both Mike and Fever _____ traveling as fast as _____ to avoid getting drenched _____ the rain. Fever, jet _____ and fearless, had been _____ horse for as long _____ they could remember. Suddenly, _____ reared back coming to _____ complete stop. Fever and _____ stood shivering in front _____ a tremendous fallen tree.

_____ hesitated, he knew Mike _____ him to jump over _____ tree and continue on _____ journey, but something was _____ him from galloping over _____. The horse was afraid, _____ even he didn't know _____ was scaring him.

When _____ was a young colt _____ attempted to jump over _____ tree that had fallen. _____ was not successful, and _____ fractured his leg.

Test 5

The queen was disturbed. Her pet owl, Wise, _____ not feeling well at _____. His feathers dropped and _____ "whoooo" was exceptionally weak. _____ had been sick before, _____ he was never this _____. He wouldn't even spread _____ wings for a special _____ treat, even his favorite _____.

The queen decided to _____ the doctor. She wouldn't _____ by and watch Wise's _____ get droopier and droopier.

_____ the time the doctor _____ Wise was too weak _____ stand on his perch. _____ a careful examination the _____ said that Wise had _____ very unusual sickness. The _____ way for him to _____ better would be to _____ leaves from the Binga-Banga _____. There was only one _____ tree in the whole world.

Test 6

David and Donnie had been best friends for as far back as either of them could remember. They shared a huge _____ story tree house in _____ giant tree that divided _____ backyards of the two _____.

The tree house was _____ around town as David _____ Donnie's Detective Den. Ever _____ David and Donnie solved _____ haunted house mystery they _____ considered two of the _____ and most successful detectives _____ the city.

When the _____ for the sixth grade _____ was discovered missing, David _____ Donnie were the two _____ people to turn to _____ help. When Lisa went _____ count the money on _____ afternoon the box was _____. The missing sixth grade _____ was a problem in _____, but eighty-five dollars was _____ reported stolen from the _____.

Test 7

The wizard drew his black cape close to his body to keep out the cold. The forest seemed to _____ in around him as _____ temperature dropped. He was _____ the center of the _____ Forest. The queen of _____ green elves was meeting _____ there at exactly midnight.

_____ was not well in _____ land. The wizard had _____ danger for weeks now, _____ until the queen of _____ elves had called this _____ he was uncertain of _____ extent of the trouble. _____ the queen's note had _____ was that the evil _____ of the West had _____ gathering troops. The wizard _____ this meant an attack. _____ wizard sighed. He was _____ now, he doubted the _____ of his power.

APPENDIX C

HOMAN MULTIPLE CHOICE CLOZE PASSAGES

HOMAN MULTIPLE CHOICE CLOZE

Test 1

Look at the cat. The cat is very fat. The children liked the [beautiful
fat
looked] cat. The fat cat [at
bring
ran] away. The children looked [for
was
either] the fat cat. The [down
the
fat] cat looked for the [father
now
children]. The fat cat was [lost
cat
gave]. Did you see a [high
fat
the] cat?

The fat cat [keep
came
the] back. He looked for [the
near
ran] children. He looked this [looked
way
nice]. He looked that way. [Fat
The
Put] children looked too. They [face
looked
fat] for the fat cat. [The
Several
Fat] fat cat said, "Help!" [So
Not
The] children ran. They ran [fat
than
to] the fat cat. He [for
was
touch] not lost now. The [fat
children
well] cat was home. He [able
fat
was] happy.

Test 2

Once upon a time a king, a queen, and their son Todd lived on the top of a huge mountain. They lived in a [bad dog royal] castle.

There was something [and blanket Todd] wanted very much. He [wanted now call] a puppy. He knew [clock it next] would be fun to [a play dinner] with a dog.

One [eye really day] Todd asked the king [king and for] queen for a dog. [If The Was] king thought about it [queen for in] a long time, but [its a he] wouldn't give Todd a [a direct light] answer. This made Todd [unhappy middle made].

Several days passed. Todd [Todd often of] thought about a puppy. [The Large Put] next day a large [for present car] drove up to the [river were castle] entrance. The driver walked [up it several] to the door.

"I [have
car
smile] something for Todd," he [said
up
swallow]. The man
took a [crate
passed
thought] out of the car. [Turtle
Surprise
Up] ! There were
three happy [little
the
whole] puppies in the box. [Young
The
Now] Todd was
really happy.

Test 3

Arnold was not an ordinary cat. His magic powers made

[attention
him
from] very special. It was [easy
was
black] to spot Arnold in

[any
going
careful] crowd of cats. He [close
something
was] the one who winked

[dinner
his
unusual] eye and stuck his [tail
to
something] in his mouth. If

[flag
Arnold
and] winked with his left [grandfather
in
eye] and had his tail

[if
in
tail] his mouth it meant [jacket
was
something] unusual was going to

[occur
Arnold
lick] , however, when Arnold winked [near
with
meant] his right eye

and [kept
however
notice] his tail in his [pasture
meant
mouth] it meant something else.

[It
Was
Present] meant something strange, possibly [even
his
round] frightening.

was about to [sew
tail
happen] .

One day as I [was
smile
unusual] walking down the street [I
with
sugar]

saw Arnold step out [powers
from
this] between two garbage cans. [An
I
Turtle]

decided to follow Arnold [powers
to
while] see if he was [planning
tail
yell]

on using his powers.

Test 4

It was a cold, damp, and rainy day. Both Mike and Fever

[ache
their
were] traveling as fast as [possible
ax
even] to avoid getting
drenched [by
something
but] the rain. Fever, jet [clean
him
black] and fearless,
had been [Mike's
and
dart] horse for as long [if
as
back] they could remember.
Suddenly, [Fever
by
holiday] reared back coming to [nor
was
a] complete stop.
Fever and [the
Mike
outdoors] stood shivering in front [tree
put
of] a
tremendous fallen tree.

[To
Rat
Fever] hesitated, he knew Mike [the
wanted
rob] him to jump over
[the
Fever
to] tree and continue on [spinach
their
even] journey, but something
was [cover
and
preventing] him from galloping over [it
even
sunshine]. The
horse was afraid, [but
so
were] even he didn't know [Fever
what
tired] was
scaring him.

When [had
Fever
valley] was a young colt [he
and
we] attempted to jump
over [tree
a
with] tree that had fallen. [He
Shivering
Youth] was not
successful, and [he
a
accident] fractured his leg.

Test 5

The queen was disturbed. Her pet owl, Wise, baking
owl
was
 not feeling well at all
disturbed
but. His feathers dropped and
his
the
cap "whoooo" was exceptionally weak. Cocoon
By
Wise had been sick
 before, but
down
Wise he was never this elbow
sick
that. He wouldn't even
 spread his
never
flag wings for a special goat
too
owl treat, even his
 favorite but
kind
wouldn't.

The queen decided to insult
wings
call the doctor. She wouldn't
lend
a
sit by and watch Wise's metal
feathers
had get droopier and droopier.
By
Of
Sick the time the doctor piled
arrived
and Wise was too weak
rather
watch
to stand on his perch. After
Sit
Same a careful examination
 the doctor
to
silk said that Wise had such
a
his very unusual sickness.
 The best
tame
on way for him to treat
get
his better would be to weed
chew
feathers

leaves from the Binga-Bange [accident
tree
the] . There was only one

[such
but
wise] tree in the whole world.

Test 6

David and Donnie had been best friends for as far back as either of them could remember. They shared a huge

[two divided barely] story tree house in [but giant the] tree that divided [the were down] backyards of the two [cod that houses] .

The tree house was [deal grade known] around town as David

[either and two] Donnie's Detective Den. Ever [from since discovered] David and

Donnie solved [in missing the] haunted house mystery they [held detective were]

considered two of the [best of instant] and most successful detectives

[in solved of] the city.

When the [money considered metal] for the sixth grade [oatmeal party most]

was discovered missing, David [put help and] Donnie were the two

[logical money rainy] people to turn to [for Lisa same] help. When Lisa went

[to missing so] count the money on [silk ever Friday] afternoon the box was

[squash
tree
empty] . The missing sixth grade [tank
money
when] was a problem in

[itself
and
train] , but eighty-five dollars was [while
also
giant] reported stolen

from the [office
shared
afternoon] .

Test 7

The wizard drew his black cape close to his body to

keep out the cold. The forest seemed to [bat
elves
close] in around
him as [had
the
but] temperature dropped. He was [at
forest
down] the center
of the [Dark
of
cranberry] Forest. The queen of [the
elves
each] green elves
was meeting [the
him
downstairs] there at exactly midnight.

[From
Was
All] was not well in [gallon
their
the] land. The wizard had
[him
sensed
has] danger for weeks now, [but
close
if] until the queen of
[been
on
the] elves had called this [meeting
the
mice] he was uncertain of
[out
of
the] extent of the trouble. [Put
Weeks
All] the queen's note had
[midnight
indicated
receive] was that the evil [seesaw
was
prince] of the West had [smack
his
been]
gathering troops. The wizard [wizard
suspected
stung] this meant an attack.
[This
The
Black] wizard sighed. He was [warm
the
older] now, he doubted the
[strength
yard
of] of his power.

APPENDIX D

KIDDER MULTIPLE CHOICE CLOZE PASSAGES

KIDDER MULTIPLE CHOICE CLOZE

Test 1

People. People. Many, many people. [Afternoon
People
The] are
in the houses. [Elephant
Many
Houses] are in the city. [Streets
Men
In] are
in the city. [People
In
Sled] are in the city.

In the city, lights go on. Boys and girls are [in
can
above]
the houses. The lights [burn
the
go] out. "Good night."
"Good [night
now
eye] ."

Dick said, "Look at Puff. Puff wants a ride.

[Many
Sally
Hen] said, "No, no, Puff! [Meat
You
The] can not ride in
[go
the
put] car. Go away, Puff. [I
Know
Good] want Tim to ride [now
true
in] ."

Debbie said, "Come here, David. I will help you

[I
bark
clean] your shoes."

"Let me [David
flew
do] it," said David. You [nod
it
can] clean your

shoes. But [you
here
store] do not have to [clean
whisper
I] my shoes. I do

[act
not
me] want help."

Test 2

White Cloud lay on his blanket and looked at the dark sky. He heard the other [to children airplane] as they were going [maybe how home] to sleep. He could [hear Billy's come] the mothers and fathers [as either hear] they talked. At last [are of the] night was quiet.

Dick said, "Look at this little red boot. Is it Billy's?"

"No, [that next mothers] isn't Billy's boot," said [will peanut Pete] .

"His boots are black."

" [I Were Ring] guess someone in our [neighborhood on ship] lost it,"

said Dick.

" [Story I'll Isn't] take the boot to [and my think] house. Maybe someone will [was come I] along and ask about [also it looked] ."

Test 3

The wind was growing stronger, and it was getting

harder and harder to walk. In his hand, the

| |
|------|
| air |
| boy |
| stay |

 tightly held the book

| |
|--------|
| caused |
| that |
| but |

 caused him to stay

| |
|-----------|
| out |
| dangerous |
| so |

 so late. Eleven-year-

| |
|------|
| old |
| so |
| fine |

 Abe could not lose

| |
|------|
| old |
| if |
| this |

 book. He had walked

| |
|--------|
| mean |
| more |
| borrow |

 than three miles to

| |
|----------|
| peek |
| fiercely |
| borrow |

 it. The wind was

| |
|---------|
| so |
| blowing |
| sand |

 so fiercely now that

| |
|-------|
| the |
| store |
| it |

 almost snatched the book

| |
|-------|
| miles |
| from |
| under |

 his hand.

Test 4

You sit in your seat at the circus, waiting for the

show to begin. You are one among [thousands
sit
afternoon] of other eager,

waiting [of
branch
people]. Perhaps you are sitting [play
crazy
high] up, near

the roof [of
are
from] the huge tent -- the [High
Up
Big] Top itself. Or

you [love
you
may] be seated lower down [pink
the
close] to the sawdust ring.

[To
Shady
Now] the band begins to [stood
play
thousands]. Somewhere far

off a [village
lion
down] roars.

Test 5

Freddie Applegate loved the sea beside his town. He

liked to fish time
and
from his rowboat, and he liked
burn
the to wade
among the daughter
among
rocks by the shore and search
he
forget for clams and
crayfish.

Horseshoe
Freddie
Liked also liked to know about
he
next the history of
his plan
to
town, because everywhere he looked so
there
rocks were things
to show history
among
that it had been there the
a
liked long long time.

Test 6

Once your dog is able to wear a collar, you must

introduce him to the leash. First let him smell [about
it
touch] ,
touch it, and lick [it
able
below] so that he can [other
crush
see] that a leash
is [and
harmless
finger] and won't bite or [hid
scratch
house] him. Let him run
[around
him
near] the house with the [can
leash
peace] dragging from his collar.
[When
Second
Loses] he loses all fear [such
of
can] a leash, you can [take
wait
it] up
the other end [among
dragging
and] begin guiding him.

Test 7

Grendel came one midnight to the hall. Tall and broad

he [was
those
air] with a skin of [horny
Heorot
bold] green which no sword

[might
midnight
click] pierce. He strode towards [dresser
open
Heorot] through the

darkness. He [thrust
horny
fried] open the door of [no
he
the] hall and found

within [hall
a
near] bank of warriors fast [oak
huge
asleep]. In his huge

clawed [hands
snatched
ponies] he snatched up thirty [some
broad
of] those unhappy

men and [bore
tall
stand] them off to his [uniform
the
home] in the fens.

APPENDIX E
INFORMAL READING INVENTORY

ORAL READING SELECTION 1
(79 Words; First Grade Level)

Fun With Our Friends
(pp. 121-122)

Preparation

This is a story about a little girl on a farm. As you read aloud find out why her Grandparents became worried.

Selection

"Sally! Sally!" called Grandmother. "Come here, Sally." Jane said, "Sally went for a walk with Tim. I guess she went to see the chickens. She wanted to see them eat." Dick ran to the chicken house to get Sally. He saw big hens in there, but he did not see Sally. "Grandmother!" he called. "Sally is not here. She is not in the chicken house." "Oh dear!" said Grandmother. "Where did Sally go? We will have to find her."

Comprehension Questions

- _____ 1. Why were Sally's Grandparents worried about her?
(Because Sally was not where they thought she would be /or/ Sally was not in the chicken house)
- _____ 2. Who told Grandmother where Sally went?
(Jane)
- _____ 3. What word in the story means "to decide without being sure?"
(to guess)

- _____ 4. Where did Dick run to get Sally?
(the chicken house)
- _____ 5. Read the sentences which show worry about Sally.
("Oh dear!" said Grandfather. "Where did Sally go?")

ORAL READING SELECTION 2
(110 Words; Second Grade Level)

Friends Old and New
(pp. 224-225)

Preparation

This is a story about two cats. As you read aloud find out which one was scared.

Selection

One day old Merry Cat took Tiny Cat far into a cornfield to catch a mouse. The corn was all picked, so there were no crows around. But a large scarecrow was still in the field. It was dressed in a man's clothes, and around its neck was a big red handkerchief. Tiny Cat said, "That old scarecrow scares me. It does! It does!" "I think it's funny," said Merry Cat. "Farmer Gray made it with two sticks. You saw him do it yourself." Tiny Cat said, "Yes, I know I did. The scarecrow didn't scare me then, but it does scare me now. I can see the handkerchief moving!"

Comprehension Questions

- _____ 1. Which cat became scared?
(Tiny Cat)
- _____ 2. Why did Merry Cat and Tiny Cat go out into the cornfield?
(to catch a mouse)

- _____ 3. Read that part of the sentence which tells us
the time of the year.
(the corn was all picked)
- _____ 4. What did the two cats see in the cornfield?
(a scarecrow)
- _____ 5. How did the scarecrow scare Tiny Cat?
(it saw the handkerchief move)

ORAL READING SELECTION 3
(106 Words; Third Grade Level)

More Friends Old and New
(p. 149)

Preparation

This is a story about a firehouse cat. As you read aloud find out what happened when he fell asleep on the firetruck.

Selection

The huge ladder truck soon came to a squealing stop in front of a burning store. The hose truck pulled up just ahead of it. Firemen hopped off the trucks and began pulling at the hoses. A large crowd of people was standing nearby, watching the bright fire. But the crowd moved back as Fireman Pat led the way to the burning store. In a few minutes water came rushing out of the hoses. It hit the fire with a loud roar. The bright fire and the roaring noise scared Snowball. He curled up under the long ladders on the truck, trying his best to hide.

Comprehension Questions

- _____ 1. What happened to the firehouse cat when he fell asleep on the firetruck?
(he went on a fire-call)
- _____ 2. What was burning?
(a store)

- _____ 3. What kind of fire truck did the cat ride?
(ladder truck)
- _____ 4. Read the sentence which tells what the cat did
when he became scared.
(he curled up under the long ladders on the
truck)
- _____ 5. Find the word that means the same as "close."
(nearby)

ORAL READING SELECTION 4
(91 Words; Fourth Grade Level)

Roads to Follow
(pp. 154-155)

Preparation

This is a selection about the Pilgrims. As you read aloud find out why the Pilgrims left their country.

Selection

William Bradford, who was governor of Plymouth longer than any other man, wrote a history of the tiny settlement. In his history he told how plans were made to come to America. He explained why he and the other men and women decided to leave their own country. Governor Bradford called these people "pilgrims" because they set out to find a land where they could be free to worship in the way they wished. Many years later everyone who talked or wrote about this small company of settlers called them Pilgrims.

Comprehension Questions

- _____ 1. Why did the Pilgrims leave their country for a strange land?

(so they could be free to worship in the way they wished)

- _____ 2. Who was the first to call these people Pilgrims?
(William Bradford)

- _____ 3. Who was governor of Plymouth longer than any other man?

(William Bradford)

- _____ 4. Name the word in the story which means "a small village."

(settlement)

- _____ 5. Find the word that means "to express a feeling of respect for God."

(worship)

ORAL READING SELECTION 5
(105 Words; Fifth Grade Level)

Ventures
(p. 225)

Preparation

This is a selection about the Monarch butterfly. As you read aloud think of a good name for this article.

Selection

East of the Rockies, as far as anyone now knows, there is no one spot where so many butterflies gather. But the flight of the eastern monarchs is amazing. Butterflies from southern Canada and northern United States fly all the way to Mexico and the tip of Florida - a great distance for creatures with four-inch wings. There are many questions about the flight of the monarchs that scientists are trying to answer. Why do the butterflies go to the same trees each year? How do they find their way over such great distances? What routes do they take, and how far south do they go?

Comprehension Questions

- _____ 1. What would be a good name for this article?
(the Mysterious Monarch /or/ the Amazing Monarch)
- _____ 2. To which two places do the monarchs from Canada and northern United States fly?
(Mexico and tip of Florida)

- _____ 3. Why are the monarchs considered mysterious?
(there are many questions about their flight
which have not been answered)
- _____ 4. Find the word that means the same as "a living
being, animal or human."
(creature)
- _____ 5. How long are the wings of the monarch butterfly?
(four inches)

ORAL READING SELECTION 6
(80 Words; Sixth Grade Level)

Vistas
(p. 207)

Preparation

This selection is about a famous doctor and his animal friends. As you read orally see if you can name in which part of the world the story takes place.

Selection

All the sounds of the Ogowe country are wild and fearful sounds - the trumpeting of elephants, the screams of panthers, the cries of gorillas, the stomping of buffaloes. And the silences are wild and fearful, too. A crocodile, pushing his eyes up over the edge of the water, is silent. A white-bellied shark, streaking upriver, is as silent as a pointed knife. Up and down the Ogowe River there is wildness like the wildness when the world first began.

Comprehension Questions

- _____ 1. On which continent do you believe this story takes place?

(Africa)

- _____ 2. What does "and the silences are wild and fearful too" mean?

(though silent at times in this wilderness there is danger everywhere)

- _____ 3. Why is the white bellied shark as silent as a pointed knife?
(because it can hurt without making a sound)
- _____ 4. Find the word that means "a heavy and lively step."
(stomping)
- _____ 5. Name four forest creatures mentioned in the story.
(accept any four of the following - elephants, panthers, gorillas, buffaloes, crocodile, shark)

ORAL READING SELECTION 7
(100 Words; Seventh Grade Level)

Vistas
(p. 282)

Preparation

This selection is about an unusual girl from Oregon who had a strong desire to study medicine. As you read orally determine her character.

Selection

When Bethenia left Roseburg that stormy night, she hoped to enter Jefferson College in Philadelphia, one of the best medical schools in America. However, when she got there and applied for admission, the professor who interviewed her only laughed. He said that Jefferson did not admit women and was not about to admit them. In fact, the Jefferson students had recently hurled rotten eggs at the female students of a medical school that did admit women. The professor did not say he approved of such conduct, but he let Bethenia clearly understand that neither did he approve of female doctors.

Comprehension Questions

- _____ 1. How would you characterize Bethenia?
(optimistic and determined)
- _____ 2. Why did the Jefferson students hurl rotten eggs at female students of a medical school?
(the students were reflecting the attitude of their professors)

- _____ 3. How did the interviewing professor first show his prejudice against female medical students?
(when he laughed at her request for admission)
- _____ 4. What was the reputation of Jefferson College?
(one of the best medical schools in the country)
- _____ 5. Why did Bethenia want to become a medical doctor?
(she saw a need for female doctors /or/ she believed women could be just as successful as men in the medical profession)

ORAL READING SELECTION 8
(167 Words; Eighth Grade Level)

Challenges
(pp. 438-440)

Preparation

This selection is about one of the great scientists of the twentieth century, Dr. Albert Einstein. After retirement from the Institute for Advance Study at Princeton he became depressed over his role in the development of the atom bomb. However, his attitude changed after being cheered by students while crossing the Princeton campus. As you read orally note his change in attitude toward atomic energy.

Selection

Still smiling, he rounded the campus on his way home. He realized he felt at least ten years younger. He had decided that his usefulness to the world was not a thing of the past. There was much he could still do. "The very menace of nuclear warfare may intimidate the human race into bringing order into its international affairs which, without the pressure of fear, it would not do," he dictated that very night. Within the next few days he had determined to overcome his dislike of all public affairs and community projects and had agreed to become the chairman of the Emergency Committee of Atomic Scientists with the idea of finding and developing peacetime uses for atomic energy so that it could be used

for "the benefit of mankind." He began writing articles for magazines and making commitments to speak on radio and television concerning the infinite number of ways in which people in every part of the world could be benefited by atomic energy.

Comprehension Questions

- _____ 1. How did Professor Einstein's attitude toward atomic energy change?
(he concluded that people in every part of the world could benefit from peaceful uses of atomic energy)
- _____ 2. How did Professor Einstein attempt to justify the atomic bomb?
(the threat of atomic warfare may compel nations to forego warfare in resolving differences among and between nations)
- _____ 3. What did Professor Einstein determine to do shortly after the campus episode?
(to overcome his dislike of all public affairs and community projects)
- _____ 4. Of what important organization did Professor Einstein become chairman?
(Emergency Committee of Atomic Scientists /or/ Committee of Atomic Scientists)
- _____ 5. What is the meaning of the word "infinite" in the phrase "infinite number of ways"?
(endless /or/ limitless /or/ unending)

LIST OF REFERENCES

- Alexander, H. W. An investigation of the cloze procedure as a measuring device designed to identify the independent, instruction, and frustration reading levels of pupils in the intermediate grades (Doctoral dissertation, University of Illinois, 1968). Dissertation Abstracts, 1969, 29, 4314. (University Microfilms No. 69-10, 625)
- Anderson, T. Cloze measures as indices of achievement comprehension when learning from extended prose. Journal of Educational Measurement, 1974, 2, 83-92.
- Baker, F. B., & Martin, T. Fortap: A Fortran test analysis package, occasional paper no. 10. Office of Research Consultation, College of Education, Michigan State University, 1970.
- Betts, E. A. Foundations of reading instruction with emphasis on differentiated guidance. New York: American Book Co., 1950.
- Bloomer, R. H., Louthan, V. H., & Read, M. Non-overt reinforced cloze procedure. U. S. Office of Education Project Report No. 2245, University of Connecticut, 1966.
- Bormuth, J. R. Comparable cloze and multiple choice comprehension test scores. Journal of Reading Behavior, 1967, 10, 291-299.
- Bormuth, J. R. Cloze test reliability: Criterion reference scores. Journal of Educational Measurement, 1968, 5, 189-196.
- Bormuth, J. R. Factor validity of cloze tests as measures of reading comprehension ability. Reading Research Quarterly, 1969, 4, 358-365.
- Campbell, D. I., & Fisk, D. W. Convergent and discriminant validation of the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Carroll, J. Defining language comprehension: Some speculations. In R. Freedle and J. Carroll (Eds.), Language comprehension and the acquisition of language. New York: Wiley, 1972.

- Coleman, E. B., & Miller, G. R. A measure of information gained during prose learning. Reading Research Quarterly, 1968, 3, 369-386.
- Cranney, A. G. A comparison of a free response and multiple choice cloze reading test (Doctoral dissertation, University of Minnesota, 1968). Dissertation Abstracts, 1969, 29, 811. (University Microfilms No. 68-12, 251)
- Cranney, A. G. The construction of two types of cloze reading tests for college students. Journal of Reading Behavior, 1973, 5, 60-64.
- Dale, E., & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 11-20.
- Dixon, W. J., & Brown, M. B. (Eds.). BMDP-77, Biomedical Computer Programs, P-Series. Berkeley: University of California Press, 1977.
- Dunkeld, C. G. The validity of the designation of instructional reading levels: A study of the relationships between children's gains in reading achievement and the difficulty of instructional materials (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1970). Dissertation Abstracts International, 1971, 31, 6274. (University Microfilms No. 71-14733)
- Fillenbaum, S., Jones, L. V., & Rappoport, A. The predictability of words and their grammatical class as a function of rate of deletion from a speech transcript. Journal of Verbal Learning and Verbal Behavior, 1964, 2, 186-194.
- Flesch, R. A new readability yardstick. Journal of Applied Psychology, 1948, 32, 221-233.
- Greene, F. P. Modification of the cloze procedure and changes in reading test performance. Journal of Educational Measurement, 1965, 2, 213-217.
- Gullford, J. P., & Fruchter, B. Fundamental statistics in psychology and education. New York: McGraw-Hill, Inc., 1973.
- Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, 1973, 65, 294-300.
- Guthrie, J. T., Seifert, M., Burnham, N., & Caplan, R. The maze technique to assess, monitor reading comprehension. The Reading Teacher, Nov., 1974, 161-168.

- Hambleton, R. K., & Cook, L. Latent trait models and their use in the analysis of educational test data. Journal of Educational Measurement, 1977, 2, 75-91.
- Harris, A. J., & Sipay, E. R. How to increase reading ability (6th ed.). New York: David McKay Co., Inc., 1975.
- Havlicek, P. Effect of violation of assumptions upon significance levels of the Pearson r. Psychological Bulletin, 1977, 84, 373-377.
- Hays, W. L. Statistics for the social sciences (2nd ed.). New York: Holt, Rinehart and Winston, Inc., 1973.
- Horton, R. J. The construct validity of cloze procedure: An exploratory factor analysis of cloze, paragraph reading, and structure of intellect tests. Reading Research Quarterly, 1974-75, 10, 248-251. (Abstract)
- Hotelling, H. The selection of variates for use in prediction, with some comments on the general problem of nuisance parameters. Annals of Mathematical Statistics, 1940, 11, 271-283.
- Jongsma, E. The cloze procedure as a teaching technique. Newark, Del.: Eric/Crier and the International Reading Association, 1971.
- Kidder, S. Comparing Rasch calibration of passage reading difficulty with Spache or Dale-Chall passage readability estimates. Paper presented at the American Educational Research Association Conference, New York, April, 1977.
- Killgallon, P. A. A study of relationships among certain pupil adjustments in language situations. Unpublished doctoral dissertation, Pennsylvania State College, 1942.
- Kirk, R. Experimental design: Procedures for the behavioral sciences. Belmont, Calif.: Brooks/Cole Publishing Co., 1968.
- Louthan, V. Some systematic grammatical deletions and their effects on reading comprehension. English Journal, 1965, 54, 295-299.
- MacGinitie, W. H. Contextual constraint in English prose paragraphs. Journal of Psychology, 1961, 51, 121-130.
- Millsap, L. A study of teachers' awareness of frustration reading level among their pupils in basal readers (Doctoral dissertation, University of Oregon, 1962). Dissertation Abstracts, 1963, 23, 2809. (University Microfilms No. 63-01, 081)

- Nie, N. H., Hull, C. H., Jenkins, J., Steinbrenner, K., & Brent, D. H. SPSS: Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1970.
- Ohmacht, F. W., Weaver, W. W., & Kohler, E. T. Cloze and closure: A factorial study. Journal of Psychology, 1970, 74, 205-217.
- Oliver, M. E. Organizing for reading instruction. Elementary School Journal, 1970, 71, 97-104.
- O'Reilly, R. P., & Streeter, R. E. Report on the development and validation of a system for measuring literal comprehension in a multiple choice format. Journal of Reading Behavior, 1977, 9(1), 45-69.
- Patty, D. Evaluating children's reading abilities. In L. Baryney (Ed.), Building reading power. Dubuque, Iowa: Kendall/Hunt Publishing Co., 1969.
- Porter, D. Modified cloze procedure: A more valid reading comprehension test. English Language Teaching, 1976, 30(2), 151-155.
- Powell, W. R. The informal reading inventory. Unpublished monograph, Center for Reading Research and Instruction, University of Illinois, 1969.
- Powell, W. R. Measuring reading performance informally. Paper presented at the International Reading Association Pre-Convention Institute on Perspectives on Testing, Houston, 1978.
- Rankin, E. F. An evaluation of the cloze procedure as a technique for measuring reading comprehension (Doctoral dissertation, University of Michigan, 1958). Dissertation Abstracts, 1958, 19, 733. (University Microfilms No. 58-03722)
- Rankin, E. F. The cloze procedure - its validity and utility. In O. S. Causey and W. Eller (Eds.), Eighth yearbook of the National Reading Conference, 1959, 3, 131-144.
- Rankin, E. F. The cloze procedure - a survey of research. In E. L. Thurstone and L. E. Hafner (Eds.), The philosophical and sociological bases of reading, National Reading Conference Yearbook, 1965, 133-150.
- Rankin, E. F. The cloze procedure revisited. In P. L. Nacker (Ed.), Interaction: Research and practice for college-adult reading, Twenty-third Yearbook of the National Reading Conference, 1974.

- Rankin, E. F., & Culhane, J. W. Comparable cloze and multiple choice comprehension test scores. Journal of Reading, 1969, 13, 193-198.
- Rufener, J. B. Use of the cloze procedure with Thai school children: An exploratory study of readability and individual differences in reading (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1972). Dissertation Abstracts International, 1972, 33, 2774. (University Microfilms No. 72-19, 919)
- Salzinger, K., Portnoy, S., & Feldman, R. The effect of order of approximation to the statistical structure of English on the emission of verbal responses. Journal of Experimental Psychology, 1962, 64, 52-57.
- Sipay, E. A comparison of standardized reading achievement test scores and functional reading levels (Doctoral dissertation, University of Connecticut, 1961). Dissertation Abstracts, 1961, 22, 2639. (University Microfilms No. 61-05431)
- Spache, G. A new readability formula for primary grade reading materials. Elementary School Journal, 1953, 53, 410-413.
- Spache, G. Good reading for poor readers. Champagne, Ill.: Garrard Press Publishing Co., 1977.
- Taylor, W. Cloze procedure - a new technique for measuring readability. Journalism Quarterly, 1953, 33, 123-145.
- Taylor, W. Recent developments in the use of cloze procedure. Journalism Quarterly, 1956, 33, 42-46.
- Taylor, W. Cloze readability scores as indices of individual differences in comprehension and aptitude. Journal of Applied Psychology, 1957, 41, 19-26.
- Weaver, W. W., & Kingston, A. J. A factor analysis of the cloze procedure and other measures of reading and language ability. Journal of Communications, 1963, 13, 252-261.
- Williams, J. L. A comparison of standardized reading test scores and informal reading inventory scores. Unpublished doctoral dissertation, Southern Illinois University, 1963.
- Wright, B. D., & Panchapakesan, N. A. Procedure for sample-free item analysis. Educational and Psychological Measurement, 1969, 29, 23-48.

BIOGRAPHICAL SKETCH

Susan Homan attended Union Avenue Elementary School in Irvington, New Jersey, through the seventh grade. She and her family moved to Hollywood, Florida, in 1960. Her schooling continued there at McNicol Junior High School and later at South Broward High School.

After graduating from high school in 1965, Susan attended the University of Florida, studying elementary education. She graduated in March, 1969, and taught kindergarten for four months at Lake Butler Elementary School in Union County, Florida.

For the next two years she taught first graders in St. Cloud, Florida. She moved to Mississippi in 1971, and became a Title I reading teacher. During her three years in Water Valley, Mississippi, as a Title I teacher, she entered the University of Mississippi, and began work on a master's degree in reading.

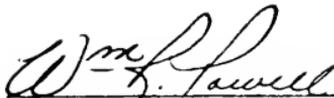
In 1974, she transferred to Florida Technological University where she completed her master's degree in reading in August, 1975. During the 1974-75 school year she taught in St. Cloud, Florida, at the fourth grade level. She was nominated by her principal for Teacher of the Year.

In September of 1975, she entered the Ph.D. program in curriculum and instruction at the University of Florida.

After completing her course work in August, 1977, she began teaching reading education courses at Florida Southern College in Lakeland, Florida.

Susan is currently teaching reading education courses as an Assistant Professor at the University of South Florida in Tampa, Florida.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



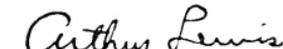
William R. Powell, Chairman
Professor of Curriculum and
Instruction

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Linda Crocker
Associate Professor of
Educational Foundations

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Arthur Lewis
Professor of Curriculum and
Instruction

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Lawrence L. Smith

Lawrence L. Smith
Associate Professor of
Curriculum and Instruction

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William B. Ware

William B. Ware
Professor of Educational
Foundations

This dissertation was submitted to the Graduate Faculty of the Division of Curriculum and Instruction in the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 1978

W. A. Sisk

Dean, Graduate School

276
F. J.
H. J. Wood
1922

EDUCATION
LIBRARY

UNIVERSITY OF FLORIDA



3 1262 08552 7884