AN INVESTIGATION INTO THE ABILITY

OF PROVIDING RATIONALE FOR CLOZE

RESPONSES AT THE THIRD AND

FIFTH GRADES

by

Hazel H. Zahradnik

A thesis

submitted in partial fulfillment

of the requirements for the degree of

Master of Education in the Department

of Curriculum, Humanities and Social Sciences

The University of Manitoba

October 1978

AN INVESTIGATION INTO THE ABILITY

OF PROVIDING RATIONALE FOR CLOZE

RESPONSES AT THE THIRD AND

FIFTH GRADES

BY

HAZEL H. ZAHRADNIK

A dissertation submitted to the Faculty of Graduate Studies of
the University of Manitoba in partial fulfillment of the requirements
of the degree of

MASTER OF EDUCATION

© 1978

ABSTRACT

This study was designed to examine, at the third and
fifth grades, the relationship between the ability to complete
cloze items with exact word replacements and the ability to
give acceptable rationale responses for cloze items. In addi-
tion, it examined the relationship between cloze scores and
measures of vocabulary, comprehension and intelligence as well
as sex.

Many researchers have shown a positive relationship
between cloze scores and measures of reading comprehension,
vocabulary, intelligence and various aptitudes, however there
is little research which investigates whether readers who score
well on cloze tests are also better able to verbalize their
reasons for their cloze responses than are the readers who score
low on cloze tests. In other words, are good readers more skill-
ed at using the syntactic, semantic and grammatical clues the
psycholinguists theorize efficient readers use? The cloze pro-
cedure was seen as a suitable tool for tapping the rationale
ability of readers so that comparison could be made between good
and poor readers.

Two cloze tests were designed, one to be used with grade
three and the other to be used with grade five. Each test was
based on a reader selection which was found to be at the grade
placement level of the pupils. An every fifth, any word deletion

pattern was applied until fifty deletions were made. Pupils were requested to complete these tests independently in writing. Following the completion of these tests, fifteen of the high cloze scores and fifteen of the low cloze scores in each grade were selected for interviews, which resulted in a total of sixty individual interviews. Pupils were requested to give reasons for all their cloze responses, however, only the rationale given for exact word replacements were evaluated and categorized into one of five possible categories: A--acceptable; R1--repetition; R2--"only one that fits"; N--not acceptable; O--omitted. Total responses in each category were computed and comparisons were made to determine whether there was a significant difference between high and low cloze scores.

Following this phase of the study, pupils were given reading tests, the <u>Gates MacGinitie Reading Tests</u>, with Primary C being administered to grade three and Survey D to grade five. Also, the <u>Otis-Lennon Mental Ability Test Form J</u> was administered, with Elementary I level given to grade three and Elementary II given to grade five.

The results of the study indicated that high scorers in vocabulary, comprehension and intelligence tests all had significantly higher total cloze scores than did low scorers in these same measures. In both grades there was no difference found between boys and girls with regard to total cloze scores.

It was also found that high cloze scorers in grade three were significantly better at providing acceptable rationale responses than were low cloze scorers, however at grade five

no significant differences in the acceptable rationale response scores were found when low and high cloze scorers were compared. It was concluded that this difference may be due to the fact that basic language skills are achieved by all pupils by the time they reach grade five, whereas the slow developing grade three may lack some of the basic language concepts. One other possible reason for the different finding in grades was that the grade five cloze test was not as difficult a test for that grade as was the grade three cloze test. Sentence length, story concepts and vocabulary were factors which contributed to test differences.

An analysis of rationale responses showed that both grade three and five pupils are able to use a wide range of clues and strategies when selecting words for cloze items. The main difference between the grades was that grade fives began to use labels such as conjunction, past tense and vowel. Pupils in both grades found nouns, pronouns, possessive pronouns, articles and some common prepositions easiest to replace by exact words and acceptable rationale was most frequently provided for nouns, pronouns, possessive pronouns with one or two adverbs, adjectives and prepositions included.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

NATURE OF THE STUDY

Reading has traditionally been conceptualized by the proponents of the look-say and phonic approaches, as a sequential processing of graphic stimuli with stress placed on first decoding one letter or one sound or one word at a time in the order in which they appeared on the page, and then putting them together to comprehend the message. The psycholinguists reject this strictly sequential model. They postulate rather that:

> ...Reading is a psycholinguistic guessing game. It involves an interaction between thought and language. Efficient reading does not result from precise perception and identification of all elements, but from skill in selecting the fewest, most productive cues necessary to produce guesses which are right the first time.[1]

These productive cues that Goodman refers to are not basically the graphic cues but the linguistic cues inherent in the material. Some of the linguistic cues relevant to the reading process include the pattern of words or syntax; inflection; intonation, as indicated by punctuation, redundancy, function words; vocabulary and semantic aspects of the surrounding text.[2]

The notion that children are not merely processing

---

[1]Kenneth S. Goodman, "Reading: A Psycholinguistic Guessing Game", Journal of the Reading Specialist, May 1967, p. 126.

[2]Delores K. Kennedy and Paul Weener, "Visual and Auditory Training with the Cloze Procedure to Improve Reading and Listening Comprehension", Reading Research Quarterly V, (Summer, 1970), p. 427.

graphic cues but are actively interpreting and interacting with linguistic cues has been substantiated by studies which have analyzed errors made during oral reading.[1,2] From these studies it has been suggested that readers sample the graphic cues with good readers being more adept than poor readers at selecting the most relevant cues. Then on the basis of these graphic cues and the reader's knowledge of the syntactic characteristics of the language, a guess is made regarding the meaning and pronounciation of the word. Then this guess is checked for semantic appropriateness in terms of the preceding and following context.

Studies of errors made during oral reading have provided some very convincing evidence to support the theory that readers use many of the linguistic cues Goodman refers to as productive.

An alternate way to investigate the kinds of cues and strategies pupils use when reading would be through the cloze procedure. The very nature of this technique suits it to the study of these searching and matching processes that psycholinguists postulate readers are engaged in when reading. Using cloze maximizes the need for pupils to use linguistic cues, since the blank provides no graphic cues to the reader.

---

[1]Kenneth S. Goodman, "A Linguistic Study of Cues and Miscues in Reading", Elementary English, Vol. 42, October 1965.

[2]Rose-Marie Weber, "A Linguistic Analysis of First-Grade Reading Errors", Reading Research Quarterly V, (Summer, 1970), p. 427.

Hafner[1] found in his study with college pupils that when doing cloze tasks, poor readers did not use immediately adjacent context cues; neither did they use widely separated sections of context. Poor readers had substantially more cloze responses which were either grammatically or semantically incongruent than did the good readers.

Very few reported studies have actually requested pupils to verbalize the kinds of cues they use when reading. Jenkinson[2] concluded from her study, in which she requested pupils to state their reasons for cloze responses, that the verbalizations distinguished the high-scoring subjects as superior in reading skills, but no further analysis of the kinds of cues these pupils used is reported on.

It would be pertinent to determine whether the effective use of linguistic cues does characterize the good reader. If so, more emphasis should be placed on promoting pupil awareness and use of these linguistic cues.

Statement of the Problem

From a review of the literature, it became apparent that there was a need to explore the types of clues and kinds of

---

[1]Lawrence E. Hafner, "Relationships of Various Measures to the Cloze", New Concepts in College-Adult Reading, ed. Eric L. Thurston and L. E. Hafner, Thirteenth Yearbook of the National Reading Conference, 1964, p. 138.

[2]Marion D. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension", (Doctoral dissertation, University of Chicago, 1957), cited by Lawrence E. Hafner, "Research For the Classroom", Journal of Reading IX, (May 1966), p. 419.

reasoning that are used by both the skilled and unskilled readers, particularly in the elementary grades. Would findings with college pupils be confirmed with grade three and grade five pupils? Would skilled readers at grade three and five be superior in their ability to complete cloze deletions and to provide rationale for their responses when compared to less skilled readers? Would the more intelligent pupils be better able to complete cloze deletions? If free to initiate their own reasons and clues, would it be found that these self-initiated clues differed significantly from those identified by Ames's graduate pupils?[1] Would fifth grade pupils be more advanced than those in grade three in the types of clues they would use? Such questions became the basis for the research study. To provide a focus, two main questions were formulated:

1. What relationship exists between the ability to complete cloze items with exact word replacements and the ability to give acceptable rationale responses for cloze responses at the third and fifth grade?

2. What relationship exists between cloze scores based on exact word replacements and intelligence, reading ability and sex at the third and fifth grade?

These questions are considered and tested under several null hypotheses.

---

[1] Wilbur S. Ames, "The Development of a Classification Scheme of Contextual Aids", The Reading Research Quarterly I, (Fall 1966), p. 57.

Hypotheses

Hypothesis 1. There is no significant difference between low and high cloze scorers at grade three regarding the frequency with which acceptable rationale responses are given for cloze items.

Hypothesis 2. There is no significant difference between low and high cloze scorers at grade five regarding the frequency with which acceptable rationale responses are given for cloze items.

Hypothesis 3. There is no significant difference in the total cloze scores of low and high vocabulary scorers in grade three.

Hypothesis 4. There is no significant difference in the total cloze scores of low and high vocabulary scorers in grade five.

Hypothesis 5. There is no significant difference in the total cloze scores of low and high comprehension scorers in grade three.

Hypothesis 6. There is no significant difference in the total cloze scores of low and high comprehension scorers in grade five.

Hypothesis 7. There is no significant difference in the total cloze scores of grade threes with high or low intelligence quotients.

Hypothesis 8. There is no significant difference in the total cloze scores of grade fives with high and low intelligence quotients.

Hypothesis 9. There is no significant difference in the total cloze scores of boys and girls in grade three.

<u>Hypothesis 10</u>. There is no significant difference in the total
cloze scores of boys and girls in grade five.

Definition of Terms

1. <u>Cloze procedure or cloze technique</u>. The term cloze
procedure is used in this study to indicate a technique in which
words are deleted from a passage and underlined blanks are sub-
stituted. The individual responding is required to write in the
exact word belonging in the blank.

2. <u>Cloze Test</u>. The cloze tests for this study are based
on passages drawn from basal readers that are rated, using the
Spache Readability Formula for grade three and the Dale-Chall
Readability Formula for grade five, at or near the grade place-
ment of subjects at the time of the testing. The first sentence
of each cloze passage was left intact, followed by the applica-
tion of an every-fifth, any-word deletion pattern until a total
of fifty words were deleted. The remainder of the sentence con-
taining the last deletion and the last sentence were left intact.

3. <u>Cloze Item</u>. A cloze item occurs when a word is de-
leted from a passage and a blank line is substituted for the
blank.

4. <u>Total Cloze Score</u>. The total cloze score refers to
the number of items for which exact word replacements were made.
The possible total cloze score is fifty for both cloze tests
used.

5. <u>Pre-Cloze Test</u>. The cloze tests in this study are
referred to as pre-cloze tests because the subjects will not

have read the original, unmutilated material prior to the taking of the cloze tests.

6. Post-Cloze Test. The cloze test is referred to as a post-cloze test when the subjects have read the original mutilated material prior to the taking of the cloze test.

7. Deletion System or Deletion Pattern. These terms are used interchangably and refer to the pre-established deletion process whereby words are removed from a printed passage. The cloze tests in this study are based on an every-fifth, any-word deletion pattern.

8. Any-Word Deletion System. This term refers to the deletion system in which every-nth word is deleted regardless of its grammatical form, as was done with the cloze tests used in this study.

9. Structural Deletion System. A structural deletion system occurs when function words such as articles, prepositions, conjunctions, and possessive pronouns are deleted on a predetermined basis.

10. Lexical Deletion System. A lexical deletion system occurs when substantive words such as nouns, verbs, adjectives and adverbs are deleted on a predetermined basis.

11. Exact Word Replacement Scoring. This scoring system refers to counting a response on a cloze test correct if it is the same word as the word deleted in the original passage, even though the spelling is incorrect.

12. Readability. This term refers to the designated level of difficulty of a particular printed material, usually

determined by the application of an accepted readability formula and expressed in grade equivalent scores.

13. <u>Acceptable Rationale</u>. Rationale responses are categorized as acceptable if the subject demonstrates the use of a clue or reasoning strategy involving: grammatical clues such as pronoun agreement, possessive agreement, or plural agreement; a syntactical or structural clue referring to word order or phrase and sentence structure; a semantical or content clue pertaining to the meaning of the passage such as main idea, supporting ideas or sequence of events.

14. <u>Repetition Rationale Response</u>. Rationale responses are categorized as repetitions when subjects repeat or paraphrase the phrase or sentence containing the cloze item, emphasizing with their voice and intonation the word placed in the blank.

15. <u>"Only One That Fits" Rationale Response</u>. Rationale responses are categorized as "only one that fits" type when subjects respond with one of several statements, for example, "You couldn't put any other word in there", "It sounded better than any other word", "What else could you put in there?", "I thought it was the only word that made sense", or words to that effect.

16. <u>Not Acceptable Rationale Response</u>. Rationale responses are categorized as not acceptable if subjects made statements demonstrating faulty reasoning or no reasoning at all.

17. <u>Omitted Rationale Response</u>. Rationale responses are categorized as omitted if subjects gave no response when rationale was requested.

Limitations

Limitations operating in this study include:

1.   This investigation is limited to analysing rationale for exact word replacements and does not include rationale for incorrect cloze items.

2.   This investigation is limited to analyzing data only for grades three and five.

3.   This investigation is limited in the number of interviews done at each grade level.

4.   The categorization of rationale responses were subject to the examiner's opinion and bias.

5.   The interview format possibly minimized some subject's ability to supply rationale since fairly spontaneous responses were required.

6.   This study is limited in that only one cloze test was used for each grade and no common cloze test was used across the grades.

Other limitations include: the size and method of selecting the sample; the faults inherent in the tests used, including the level of difficulty of the cloze tests; the incapacity of some pupils to read the cloze tests; and the contrasting formats of the intelligence tests used.

Assumptions

Several assumptions were made:

1.   That the interview format actually tapped the rationale-giving ability of subjects to their maximum.

2.   That the subjects would recall the actual process-
ing and reasoning they engaged in when doing the cloze test
independent and silently which was a day or more prior to the
actual interview.

3.   That the grade three and five classes in the study
were representative of these grades.

Overview of the Study

This study is designed to investigate the relationship
of total cloze scores to the ability to provide acceptable
rationale for cloze items at two grade levels and to investi-
gate the relationship of total cloze scores to reading ability,
intelligence and sex. Chapter 1 outlines the purpose of the
study, which is followed by a review of the literature and
research related to this research study in Chapter 2. The
design of the study is explained in Chapter 3, which includes
a description of the pilot study as well as a description of
the population, the test instruments and the research proce-
dures followed. An analysis of the data both in table and
narrative form is contained in Chapter 4. Chapter 5 consists
of a summary of the findings, conclusions drawn on the finding
and implications for classroom practice, as well as recommen-
dations for further research.

Chapter 2

REVIEW OF RELATED LITERATURE

The purpose of this study was to investigate the rela-
tionship between ability to give rationale for cloze responses
and the ability to complete cloze items. The relationship be-
tween pre-cloze scores and intelligence, vocabulary, comprehen-
sion and sex was also investigated.

Research on the cloze procedure began in 1953, when it
was first introduced by Taylor. It was originally conceived
of as a measuring instrument; thus much of the research concen-
trates on cloze as a measurement device. Many of the studies
examine the validity and reliability of cloze as a measure of
readability, reading comprehension, reading achievement and
reading gain. There have been, however, relatively few studies
investigating cloze as either an instructional or diagnostic
tool.

The literature and the research to be considered will
include cloze as a measure of readability and as a measure of
comprehension and a review of any correlation studies which
relate cloze to other measures such as reading achievement and
intelligence, or aptitude.

Studies will also be included which investigate the
various methodological aspects of constructing, scoring and
interpreting cloze tests, and any research available on the

strategies and cues subjects use when completing cloze tasks.

The Cloze Technique

When Taylor introduced the cloze procedure in 1953, he

defined it as:

> a method of intercepting a message from a 'transmitter',
> mutilating its language patterns by deleting parts, and
> so administering it to 'receivers' that their attempts to
> make the patterns whole again potentially yield a consid-
> erable number of cloze units.[1]

The term "cloze" was originally derived from the Gestalt con-

cept of "clozure", which is the tendency to complete a struc-

tured whole by filling a missing gap. According to Weaver[2]

however, the relationship of Gestalt ideas of clozure to the

cloze procedure is a tenuous one. He indicates the true Ges-

taltist has always concentrated on the perceptual and has

invariably considered clozure a literal closing. Weaver does

not think that subjects complete cloze units in this literal

way. He views cloze as a coding operation in which the subject

must engage in a searching procedure; he must survey certain

categories of language elements for the most probable one to

fit a certain context. In other words, Weaver views cloze as

a phenomenon related to the production of components to fit a

---

[1]Wilson L. Taylor, "Cloze Procedure: A New Tool of
Measuring Readability", _Journalism Quarterly_ XXX, (Fall,
1953), p. 416.

[2]Wendell W. Weaver, "Theoretical Aspects of the Cloze
Procedure", the _Philosophical and Sociological Bases of
Reading_, ed. E. L. Thurstone and L. E. Hafner, Fourteenth
Yearbook of the National Reading Conference, 1965, p. 116.

language organization. Rankin also suggested that making cloze

responses is primarily a cognitive rather than a perceptual

task.[1]

Cloze as a Measurement of Readability

Research has been carried out to support the use of

cloze as a technique for measuring readability. Taylor,[2] Bor-

muth,[3] Gallant[4] and Rankin[5] have attempted to determine the

strength of the relationship between cloze tests and the more

conventional measure of readability such as readability formu-

lae.

Taylor[6] ranked several passages by Flesch and Dale-

Chall readability formulae and by total pre-cloze, any-word

---

[1]Earl F. Rankin, "Closure and Cloze Procedure", Third
Yearbook of the North Central Reading Association, Minneapolis:
University of Minnesota, 1964, cited by Lawrence E. Hafner,
"Cloze Procedure", Journal of Reading IX (May 1966), p. 415.

[2]Wilson L. Taylor, "Cloze Procedure: A Tool for Meas-
uring Readability," pp. 416-31.

[3]John R. Bormuth, "Cloze as a Measure of Readability",
Reading as an Intellectual Activity, ed. J. A. Figurel, Newark,
Delaware: International Reading Association, 1963, p. 430.

[4]Ruth Gallant, "Use of Cloze Tests as a Measure of
Readability in the Primary Grades", Reading and Inquiry, ed.
J. A. Figurel, Newark, Delaware: International Reading Asso-
ciation, 1965, pp. 286-7.

[5]Earl F. Rankin, "Grade Level Interpretation of Cloze
Readability Scores", Reading: The Right to Participate, ed.
Frank P. Greene, Twentieth Yearbook of the National Conference
on Reading, 1970, p. 31.

[6]Wilson L. Taylor, "Cloze Procedure: A Tool for Mea-
suring Readability", p. 427.

deletion scores and he found the passages ranked the same by

each method. When he used passages by James Joyce, Gertrude

Stein and Erskine Caldwell, however, he found that the read-

ability formulae rated these as "easy" but the cloze ranking

placed them as the most difficult of all passages. He con-

cluded that cloze is able to measure relative readability on

certain types of material which cannot accurately be evaluated

by "element counting" formulae.

In a very intensive investigation of the cloze proce-

dure as a measure of readability, Bormuth ranked articles for

difficulty using cloze scores and multiple-choice criterion

tests covering the same material.[1] When he obtained correla-

tions of .92 between cloze and multiple-choice rankings, he

concluded that the readability of the cloze rankings was very

high at these three different elementary grades.

On the other hand, cloze tests were found to measure

the effect of increased sentence length on the reading diffi-

culty of passages at grade one and two. Gallant[2] compared

performances on two sets of passages rewritten as cloze tests

with the same controls on each, except one set was not con-

trolled for sentence length. After completion of her study,

Gallant concluded that cloze tests offer a means of isolating

variables which affect readability and that the cloze procedure

---

[1]John R. Bormuth, "Cloze as a Measure of Readability",
p. 430.

[2]Ruth Gallant, "Use of Cloze Tests as a Measure of
Readability in the Primary Grades", p. 287.

was a valid and reliable measure of reading comprehension for beginning readers.

Rankin's findings support the use of cloze as a means of measuring readability.[1] His results indicated a close correspondence between Bormuth's forty-four percent criterion for cloze readability analysis and the readability formulae. Therefore, Rankin concluded that the available evidence points to a high degree of validity for the cloze procedure as a measure of the relative readability of different passages for specific groups of readers.

Cloze as a Measurement of Reading Abilities

A number of studies have researched the strength of the cloze procedure as a measure of reading comprehension and reading achievement. It was concluded by Hafner that the ability to read with understanding can be measured rather well by the cloze procedure.[2] To substantiate this statement, Bormuth's study with intermediate grade students is cited in which high correlations between nine cloze tests and multiple-choice tests were obtained for seven different types of comprehension.[3]

---

[1] Earl F. Rankin, "Grade Level Interpretation of Cloze Readability", p. 427.

[2] Lawrence E. Hafner, "Implications of Cloze", Philosophical and Sociological Bases for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference, 1965, p. 151.

[3] John R. Bormuth, "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability", Reading Research Quarterly IV (Spring 1969), p. 363.

Jenkinson, in her study with high school students, obtained

correlations between cloze scores and scores in the Cooperative

Reading Test, C-2, which were .78 with vocabulary and .73 with

comprehension.[1] She concluded that the cloze procedure was a

valid measure of reading achievement.

Many other researchers have had somewhat similar results.

Rankin, using the Diagnostic Reading Test Survey as a criterion

for general reading skill obtained correlations with pre-cloze,

any-word deletion tests and pre-cloze, noun-verb deletion tests

respectively: Story Comprehension .29 and .57; Vocabulary .68 and

.42; Paragraph Comprehension .60 and .39.[2] Ruddell,[3] Fletcher[4]

and Friedman[5] all reported correlations ranging from .55 to .87

---

[1]Marion D. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension", (Doctoral dissertation, University of Chicago, 1957), cited by Earl F. Rankin, "The Cloze Procedure--A Survey of Research", Philosophical and Sociological Bases for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference, 1965, p. 135.

[2]Earl F. Rankin, "Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension", (Doctoral dissertation abstract, University of Michigan, 1958).

[3]Robert B. Ruddell, "An Investigation of the Effect of the Similarity of Oral and Written Patterns of Language Structure on Reading Comprehension", (Doctoral dissertation, Indiana University, 1963), p. 175.

[4]J. E. Fletcher, "A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities", (Doctoral dissertation, University of Washington, 1959), cited by E. F. Rankin, "The Cloze Procedure--A Survey of Research", Philosophical and Sociological Bases for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference, 1965, p. 135.

[5]M. Friedman, "The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at The University of Florida), (Doctoral dissertation, The University of Florida, 1964), cited by E. F. Rankin, "The Cloze Procedure", p. 135.

using three different standardized tests, namely the Cooperative Reading Test, Standard Achievement Test and Metropolitan Achievement Test.

Correlations when using primary children were equally high according to the Gallant study.[1] Using the primary I, II and elementary forms of the Metropolitan Achievement Tests and comparable forms written as cloze tests, she obtained correlations ranging from .65 to .81.

Using sixth grade pupils as subjects for his study, Schneyer obtained similar correlations between two cloze scores based on an every-tenth word deletion pattern and a noun-verb deletion pattern and the Gates Reading Survey.[2] Higher correlations were obtained when every tenth word was deleted as opposed to when alternating noun-verb deletions were made: vocabulary .74 and 163; comprehension .68 and .60; speed .63 and .60.

The only finding to date which conflicts with previously mentioned studies is the Weaver and Kingston finding.[3] They correlated four silently read cloze tests with the total raw

---

[1] Ruth Gallant, "Use of Cloze Tests as a Measure of Readability in Primary Grades", p. 287.

[2] J. Wesley Schneyer, "Use of the Cloze Procedure for Improving Reading Comprehension", The Reading Teacher, Vol. 19, No. 3 (December 1965), p. 177.

[3] Wendell W. Weaver and Albert J. Kingston, "A Factor Analysis of the Cloze Procedure and Other Measures in Reading and Language Ability", The Journal of Communication XII (December 1963), p. 252.

scores of the Davis Reading Test and obtained substantially lower correlations: .21 and .36 with two pre-cloze, noun-verb deletion tests and .25 and .51 with two pre-cloze, any-word deletion tests. These investigations concluded that "cloze tests are related only moderately to the verbal comprehension factor." Bormuth questioned these findings on at least four counts: selective deletion procedures were used in making some of the cloze tests; subjects were college students and likely a highly select group; the standardized tests showed unusual patterns of factor loading such as tests of rate memory of numbers, of sensitivity to grammatical structure; the cloze tests showed some inconsistencies among themselves in their loading patterns.[1] So Bormuth concluded further investigation was needed, and he proceded with a follow-up study which demonstrated that cloze tests made by deleting every fifth word measure skills closely related or identical to those measured by conventional multiple-choice reading comprehension tests. He concluded that little or no evidence was found to suggest that the two types of tests might measure different skills.

In summary, all studies except for Weaver and Kingston, have yielded high to moderately high correlations between cloze and standardized reading tests. This would seem to indicate that cloze is a valid and reliable tool for testing comprehension.

---

[1] John Bormuth, "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability," p. 361.

Methodological Considerations of Cloze

This section will review some of the research findings
on cloze test construction which relates to deletion patterns,
categories of deletion items, alternate scoring methods, test
length, age limitations, pre-cloze versus post-cloze format, and
interpretation of cloze scores.

Random Versus Every Nth Deletion

Taylor introduced the cloze procedure as using a com-
pletely 'random' counting-out system based on a table of random
numbers or simply counting out every nth word without regard
for functions or meaning of words.[1] When comparing random dele-
tion and every nth deletion, Taylor found every nth was easier
to use. On a 16-item cloze test, Taylor found a random 10% de-
letion and an every-tenth word deletion both ranked the reada-
bility of three passages the same. Most researchers have used
an every-nth procedure because it is less cumbersome than the
pure random method.

There is no rule for deciding which 'every-nth' proce-
dure, such as an every-second, every-fifth or every-tenth word,
should be used. MacGinitie studied this problem by varying the
number of words between cloze items.[2] He found that deletions
ranging from every-sixth through every twenty-fourth word were

---

[1]William L. Taylor, "Cloze Procedure: A New Tool for
Measuring Readability," p. 422.

[2]Walter H. MacGinitie, "Contextual Constraint in English
Prose Paragraphs," Journal of Psychology 51 (January 1961): 127.

almost equally restorable, however, when deletions were every

third word or less the replacement of words became much more

difficult. MacGinitie concluded that the influence of context

on the selection of a particular cloze response was highly re-

lated to the number of words between deletions. He further

concluded that every-fifth word deletions were the best possi-

ble deletion pattern.

When studies were surveyed by Rankin,[1] he found the

every-fifth word deletion pattern to be most predominant but

he concluded that this deletion pattern is not necessarily most

suitable for all materials, age groups or purposes.

It has been recommended by Culhane that an every-tenth

word count be used for textual materials that are fact-laden

and an every-fifth word count be used for narrative or literary

material.[2]

In conclusion, it seems an every-fifth, any-word deletion

pattern is most highly recommended and most frequently selected

for research studies particularly when narrative material is

being used as the basis for the cloze test.

Mechanical Versus Rational Deletions

Rankin asks the question whether it is better to con-

struct cloze tests with some element of rationality or subjec-

tivity in the selection of cloze items. He cites Green's

[1] Earl F. Rankin, "The Cloze Procedure--A Survey of Research," p. 145.

[2] Joseph W. Culhane, "Cloze Procedures and Comprehension," The Reading Teacher 23 (February 1970): 411.

study to substantiate his opinion that mechanical deletions are as effective as rational deletions.[1] Greene compared two types of deletions, the 'standard' post-cloze, any-word deletion cloze test with a 'modified' post-cloze, noun-verb-adjective deletion cloze test. The word deletions were also restricted to words which were viewed as having sufficient redundancy to permit their prediction. Greene found better reliability, better distribution of item difficulties and better item discrimination in the modified form, however, validity coefficients for the two tests were not significantly different. He concluded that the modified test did not produce a significant difference in its effectiveness in spite of all the additional time and effort needed for its construction.

To further evaluate the efficacy of subjective selection of cloze items, MacGinitie's study on 'key' word deletions is pertinent. It was initially suggested by MacGinitie[2] that if 'key' words in a selection were deleted it might reduce the subject's chances of responding correctly to other items, however, MacGinitie's later findings opposed this contention. MacGinitie found that as long as four or more words of continuous text appeared on either side of an item, the item's difficulty was unaffected by the deletion or non-deletion of other

---

[1]Frank P. Greene, "Modification of the Cloze Procedure and Changes in Reading Test Performances," _Journal of Educational Measurement_ 2 (December 1965): 216.

[2]Walter H. MacGinitie, "Contextual Constraint in English Prose Paragraphs," _Journal of Psychology_ 51 (January 1961): 128.

words in the text. Thus it can be seen that mechanical dele-
tion is the best and most efficient method of deletion that
can be used.

Types of Words to be Deleted

When constructing cloze tests, it is possible to delete
words without respect to grammatical form or to restrict dele-
tions to certain types of words. Investigations have been con-
ducted with cloze tests constructed having special types of
words deleted.

Three types of cloze tests formed by deleting: (1) any-
word; (2) easy words such as conjunctions, pronouns, articles,
verb auxiliaries; and (3) hard words such as nouns, verbs, ad-
verbs; were correlated in one of Taylor's studies.[1] Significant
correlations between all three types of cloze tests and criterion
tests of pre-reading knowledge, immediate recall and aptitude
were found. Easy-word correlations were lowest, any-word cor-
relations were highest and the hard-word form correlated best
with pre-reading knowledge. He concluded that for purposes of
testing comprehension, aptitude and readability, the any-word
form was superior.

Another hypothesis, which was formulated by Rankin[2] and

---

[1]W. L. Taylor, "Cloze Readability Scores as Indices of
Individual Differences in Comprehension and Aptitude," Journal
of Applied Psychology 41 (February 1957): 22.

[2]Earl F. Rankin, "The Cloze Procedure--Its Validity and
Utility," Eighth Yearbook of the National Reading Conference,
ed. O. S. Causcy and W. Eller, National Reading Conference, cited
by E. F. Rankin, "The Cloze Procedure--A Survey of Research,"
p. 146.

Weaver,[1] suggests that two aspects of meaning could be studied
by the cloze procedure if two different deletion procedures were
used. These two deletion procedures were classified as lexical
and structural, with lexical deletions defined as nouns and verbs
and occasionally adjectives and adverbs and structural deletions
defined as function words such as articles, prepositions, con-
junctions and possessive pronouns. Lexical deletions affect the
substantive content whereas structural deletions affect the syn-
tax or word order. Rankin tested and confirmed his prediction
that any-word deletions would best measure the group of struc-
tural meaning and noun-verb deletions best measure lexical meaning
of subjects.

Horton agreed with Rankin when he stated that the any-
word deletion procedure is not only the most efficacious way to
measure lexical words, but also results in the deletion of more
structural words.[2] He quotes Taylor as supporting the any-word
random deletion procedure because by this method there is an
assurance of a reasonable proportion of all kinds of words being
deleted in relation to their occurrence in the language. Thus
both proportion between lexical and structural words is main-
tained and the various kinds of structural words are sampled.

---

[1]Wendell W. Weaver and Albert J. Kingston, "A Factor
Analysis of the Cloze Procedure and Other Measures of Reading
and Language Ability," Journal of Communication 12 (December
1963): 253.

[2]R. J. Horton, "The Construct Validity of Cloze Pro-
cedure: An Exploratory Factor Analysis of Cloze, Paragraph
Reading and Structure of Intellect Tests," (Doctoral disser-
tation, Hofstra University, 1973), p. 125.

It would seem that there is a general agreement among researchers that the any-word deletion procedure yields the best sample of words for research purposes.

Scoring Methods

Cloze tests can be scored by the strictly objective procedure whereby credit is given only if the exact word is replaced or by a more subjective procedure in which credit is given for synonyms. Taylor tried to determine which method would be best.[1] He ranked three passages for cloze readability on the basis of objective scoring produced higher scores, but the degree of differentiation among passages was almost identical.

Rankin's findings substantiate Taylor's findings.[2] He found correlations of .86 and .92 between cloze scores using an objective and synonym scoring system for two different cloze tests and there was no significant difference in test reliability or validity for the two scoring systems.

The discrimination power of two scoring systems on reading material written with high and low frequency patterns of language structure was investigated by Ruddell.[3] He found

---

[1]W. L. Taylor, "Cloze Procedure: A New Tool for Measuring Readability," p. 425.

[2]Earl F. Rankin, "The Cloze Procedure--A Survey of Research," p. 146.

[3]Robert B. Ruddell, "A Study of the Cloze Comprehension Technique in Relation to Structurally Controlled Reading Material," Improvement of Reading Through Classroom Practice, ed. J. A. Figurel (Newark, Delaware: International Reading Association), 1964, p. 300.

that cloze tests scored by the exact word replacement method
and by the synonym count method did not differ significantly
in validity or reliability except a significantly higher reli-
ability coefficient was reported when a synonym count was used
on high frequency patterns of language structure.

A correlation of .61 was obtained by Hafner when he
compared exact word replacement scores and scores obtained by
counting responses which did not match the deleted words but
were grammatically correct.[1]

The most sophisticated study on alternate scoring was
Bormuth's study[2] in which he classified cloze test responses
into seven categories: exact word, grammatically correct;
exact word, grammatically incorrect; synonym, grammatically
correct; synonym, grammatically incorrect; unrelated semanti-
cally, grammatically correct; unrelated semantically, grammat-
ically incorrect; unclassifiable. Intercorrelations among the
cloze and reading achievement scores were found. Bormuth con-
cluded that scores obtained by counting exact responses yielded
the most valid scores and that exact response scores yielded
the greatest amount of discrimination among passage difficulties.

Several interesting scoring methods were developed by

---

[1]Lawrence E. Hafner, "Relationships of Various Measures
to the Cloze," New Concepts in College-Adult Reading, ed. Eric
L. Thurston and L. E. Hafner, Thirteenth Yearbook of the
National Reading Conference, 1964, p. 138.

[2]John R. Bormuth, "Validities of Grammatical and Seman-
tic Classifications of Cloze Test Scores," Reading and Inquiry,
ed. J. A. Figurel (Newark, Delaware: International Reading
Association), 1965, p. 285.

Hafner in his study with college students.[1] He used a ratio
of correct connective word completions such as conjunctions,
prepositions, auxiliaries and articles, to correct content
word completions such as nouns, verbs, adjectives, adverbs
and gerunds. He also used a GCIA score which he defined as
the percent of grammatically correct predictions of those
words predictions which were correct. In addition, he inves-
tigated the use of an incongruency score defined as the fre-
quency with which cloze responses were incongruent with clearly
stated information in the passage or with sentence language
patterns. All of these scoring techniques involve studying
the errors on a cloze test rather than just the correct or
acceptable responses. Further research and study is needed
before these methods can be used confidently.

The exact response scoring seems to be highly favored
as the most valid and most discriminating means of scoring
cloze test.

Length of Test

The length of a cloze test needs to be determined for
practical reasons. Rankin reports Taylor has suggested that
a fifty-item test provides a sufficient sample for a stable
score as well as chances for easy and hard words to cancel out.[1]

Further investigation by Bormuth was done to determine

---

[1]Lawrence E. Hafner, "Relationship of Various Measures
to the Cloze," p. 137.

[2]W. L. Taylor, "Recent Developments in the Use of Cloze
Procedure," Journalism Quarterly 33 (Winter 1956): 47.

how the length of the cloze test or the number of items in a
cloze test affects the difficulty levels of the different cloze
forms that are made from the same passage.[1] He made five cloze
tests of fifty items each for twenty different passages selected
from literature, history, geography and science material. The
tests were then divided after every fifth item to obtain ten
different test lengths. When subjects in grades four through
eight were given the tests, he found significant differences
among the means of the tests at all the test lengths. The dif-
ferences among the test forms that were made from the same pas-
sage tended to diminish as more items were included in the test
and the rate of reduction decreased as the number of items in-
cluded in the test became larger. Bormuth concluded that test
forms made from the same passage, if made sufficiently long,
might come to yield nearly identical mean difficulties. This
is not always practical, so Bormuth suggested researchers should
use more than one cloze test form over a passage or they should
consider calculating the standard test form error to determine
whether the mean on a single cloze test form approximates the
mean difficulty of the passage.[2]

Rankin agreed with Bormuth in that cloze tests need to
be fairly long to provide a sufficient number of discriminating

---

[1]John R. Bormuth, "Optimum Sample Size and Cloze Test
Length in Readability Measurement," Journal of Educational
Measurement 2 (June 1965): 115.

[2]John R. Bormuth, "Experimental Applications of Cloze
Tests," Improvement of Reading Through Classroom Practice, ed.
J. A. Figurel (Newark, Delaware: International Reading Asso-
ciation), 1964, p. 306.

items since mechanical selection of words for deletions tend
to produce a sizable number of non-discriminating items which
lowers test reliability.[1]

There is agreement that cloze tests need to be suffi-
ciently long to provide not only a sufficient sample but also
to approximate the mean difficulty of the passage.

Age Groups or Reading Levels

There is some controversy about the minimum age required
before a subject can function satisfactorily on a cloze test. In
1965 Rankin stated there was no systematic research on the ques-
tion. Bloomer had suggested that a pupil have at least junior
high school word recognition skills in order to make cloze res-
ponses;[2] his reason being that a pupil must understand all words
in the context of a blank space.

The great majority of studies surveyed have used high
school pupils or college pupils as subjects, however, Bormuth
used fourth, fifth and sixth grade children and obtained very
high reliability and validity coefficients.[3] Brooks reports

_____

[1] Earl F. Rankin, "The Cloze Procedure--A Survey of
Research," p. 144.

[2] R. H. Bloomer, "The Cloze Procedure as a Remedial
Reading Exercise," Journal of Developmental Reading 5 (Spring
1962): cited by Eugene Jongsma, The Cloze Procedure as a
Teaching Technique, ERIC/CRIER and the International Reading
Association, 1971, p. 7.

[3] John R. Bormuth, "Cloze Tests as Measures of Reada-
bility," (Doctoral dissertation, Indiana University, 1962),
cited by Earl F. Rankin, "The Cloze Procedure--A Survey of
Research," p. 143.

using mentally retarded elementary children[1] while Roossinck

used sixth grade children to develop programmed cloze materials[2]

and Ruddell used fourth grade pupils and found high reliability

and validity coefficients on a series of cloze tests.[3]

Since Rankin's survey, other researchers have conducted

cloze studies using primary children. Generally, they have mod-

ified the cloze task for the first grade subjects to one of

choosing between the possibility of three words.

Gallant explored the reliability and validity of cloze

tests as a measure of reading comprehension for pupils in the

first three grades.[4] She used the modified version for the

first grades in which a choice of three responses appeared for

each deleted word. Her findings indicated that the cloze tests

were both valid and reliable measures of reading comprehension

---

[1]S. T. Brooks, "Effects of Locus of Control and Anxiety
on the Ability of Mentally Retarded Children to Use Context
Clues in Reading," (Doctoral dissertation, George Peabody Col-
lege for Teachers, 1962), cited by A. C. Bickley, Billie J.
Ellington and Rachel T. Bickley, "The Cloze Procedure: A Con-
spectus," Journal of Reading 2 (Summer 1970): 238.

[2]Pearl A. Roossinck, "A Learning Program in Paragraph
Comprehension," (Master's paper, University of Michigan, 1962),
cited by Eugene Jongsma, The Cloze Procedure as a Teaching
Technique (International Reading Association, Newark, Delaware:
1971), p. 6.

[3]Robert B. Ruddell, "An Investigation of the Effects of
the Similarity of Oral and Written Patterns of Language Struc-
ture on Reading Comprehension," (Doctoral dissertation, Indiana
University, 1963), cited by Earl J. Rankin, "The Cloze Procedure
--A Survey of Research," Philosophical and Sociological Bases
for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth
Yearbook of the National Reading Conference, 1965, p. 143.

[4]Ruth Gallant, "Use of Cloze Tests as a Measure of
Readability in the Primary Grades," Reading and Inquiry, ed.
J. A. Figurel (Newark, Delaware: International Reading Asso-
ciation, 1965), p. 287.

for these primary grade children.

Using pupils representing grades one through six, Ransom compared the relative indicated reading levels of children on a cloze test and the informal reading inventory.[1] She found the correlation between these two measures was statistically significant for determining instructional and frustration reading levels for grades two through six, but not for grade one.

Pupils in grades one through six were used in Kirby's study also.[2] Her purpose was to determine whether scores derived from a cloze test would differ significantly from scores derived from selected standardized silent and oral reading tests. The Gates Reading Test, Gilmore Oral Reading Test, Gray Oral Reading Test and the cloze test were the instruments used.

Her findings indicated that the four instruments did not measure equivalently the reading achievement of pupils in grades one through four, boys, girls, low ability pupils and less able readers. The four instruments, however, did identify comparable reading for superior readers and pupils in grades five and six. She concluded that possibly the cloze test was less appropriate for young children but a promising technique for determining instructional reading levels for grades two through six.

---

[1] Peggy Ransom, "Determining Reading Levels of Elementary School Children by Cloze Testing," Forging Ahead in Reading, ed. J. A. Figurel (Newark, Delaware: International Reading Association, 1968), p. 481.

[2] Clara L. Kirby, "Using the Cloze Procedure as a Testing Technique," Reading Diagnosis and Evaluation, ed. D. L. DeBoer (Newark, Delaware: International Reading Association, 1970), pp. 75, 76.

The research of Kirchoff does not support those who
are concerned about first graders and cloze.[1] Using 120 first
graders near the end of grade one, he reported cloze was a
valuable tool in establishing instructional reading levels.
He reported that the correlations between cloze and the infor-
mal reading inventory were significant at all reading levels
in his first grade population.

In support of Kirchoff, Weaver and Kingston used the
cloze procedure as a teaching technique with culturally dis-
advantaged first graders.[2] According to these investigators,
it was demonstrated that first graders could perform on written
cloze tests soon after they began reading. They found the
multiple-choice, lexical cloze was the best single predictor
of first grade reading achievement with any-word cloze tests
and either cloze tests were better predictors than the stan-
dardized readiness test.

Given the findings of the research cited, it can be
concluded that the cloze procedure could be used with children
in all grades excepting grade one, with whom the modified cloze

---

[1]Leo H. Kirchoff, "A Study Utilizing the Cloze Test Pro-
cedure to Determine Reading Levels of First Grade Children Who
Have Been Taught Beginning Reading by Four Different Approaches,"
(Doctoral dissertation, University of Kansas, 1968), cited by
Theodore A. Mark, "Clozing the Placement Gap: A New Tool for
Administrators and Teachers," Educational Leadership 28 (April
1971): 765.

[2]A. J. Kingston and W. W. Weaver, "Feasibility of Cloze
Techniques for Teaching and Evaluating Culturally Disadvantaged
Beginning Readers," Research and Development Center, University
of Georgia, 1969, cited by A. C. Bickley, Billie J. Ellington
and Rachel T. Bickley, "The Cloze Procedure: A Conspectus,"
Journal of Reading Behavior 2 (Summer 1970): 243.

procedure would be advised.

Pre-Cloze Versus Post-Cloze

According to Rankin,[1] there is little evidence in support of either the pre-cloze format in which subjects complete the cloze test without reading the original unmutilated material prior or the post-cloze format in which subjects read the original unmutilated material before taking the test.

Taylor[2] and Rankin[3] each obtained slightly higher correlations between post-cloze scores and comprehension test scores than between pre-cloze and comprehension scores. Rankin concludes, however, that only one difference between correlations was statistically significant and furthermore, he states the extra time and effort involved in using post-cloze procedures does not yield sufficiently superior results to justify its use.

Contrasting results were obtained by Bormuth and Mac-Donald[4] when they correlated scores on a cloze test given after the subjects had studied the author's work and scores on the

---

[1] Earl J. Rankin, "An Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension," (Doctoral dissertation, University of Michigan, 1957), abstract, p. 733.

[2] W. L. Taylor, "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," p. 22.

[3] Earl J. Rankin, "An Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension," (Doctoral dissertation, University of Michigan, 1957), abstract, p. 733.

[4] J. R. Bormuth and O. L. MacDonald, "Cloze Tests as a Measure of Ability to Detect Literary Style," Reading and Inquiry, ed. J. A. Figurel (Newark, Delaware: International Reading Association), 1965, p. 287.

test of ability to detect the author's literary style. They
did not obtain significantly higher correlations using the
post-cloze procedure, however, they stated this did not neces-
sarily contradict Rankin's earlier finding. They concluded
the difference was likely due to differences in the criterion
test and the type of deletions used. Bormuth and MacDonald
conclude that cloze tests are as valid when administered to
subjects that have not read the material from which the tests
were made as when the tests were administered to subjects who
have studied those materials.

According to these findings it seems that pre-cloze
tests are as valid as post-cloze tests and in terms of prac-
ticability, pre-cloze is most efficient.

Interpreting Cloze Scores

Numerous studies have shown that cloze scores are
highly valid measures for readability of materials and for
comprehension of readers, but as Rankin and Culhane have
pointed out, there is a difficulty in interpreting raw cloze
scores.[1] What is needed is a frame of reference with which
to interpret an acceptable level of performance on cloze tests.

Bormuth attempted to provide such a frame of reference
by determining comparable scores on cloze and multiple-choice
tests.[2] In his earlier study, he used nine cloze passages with

---

[1] Earl F. Rankin and Joseph W. Culhane, "Comparable Cloze
and Multiple-Choice Comprehension Test Scores," Journal of
Reading 12 (December 1969): 197.

[2] John R. Bormuth, "Comparable Cloze and Multiple-Choice
Comprehension Test Scores," Journal of Reading 10 (February
1967): 296.

100 pupils enrolled in grades four and five. He found that a
cloze score of 38 percent was comparable to 75 percent on a
multiple-choice test and could be considered the subject's in-
structional level. He also found that a cloze score of 50
percent was comparable to a multiple-choice score of 90 percent
and could be regarded as the independent reading level of a
subject. Bormuth considered his 1968 findings more valid in
which his cloze percentages changed to 44 percent for instruc-
tional level and 57 percent for an independent level.[1]

Rankin and Culhane attempted to replicate Bormuth's
studies in order to determine a frame of reference for inter-
preting cloze scores.[2] Their findings tend to corroborate the
validity of the comparable cloze and multiple-choice scores
found by Bormuth, particularly those findings in his 1968 study.
Rankin and Culhane found 41 percent on the cloze was comparable
to 75 percent on a multiple-choice test which could be called
an instructional level, while 61 percent on the cloze was com-
parable to 90 percent on a multiple-choice test which could be
called an independent level.

Others have attempted to establish reading level cri-
teria for the cloze. Ransom constructed informal reading
inventories and cloze tests from graded basal readers and

---

[1]John R. Bormuth, "Cloze Test Readability: Criterion
Reference Scores," Journal of Educational Measurement 5
(1968): 195.

[2]Earl F. Rankin and Joseph W. Culhane, "Comparable
Cloze and Multiple-Choice Comprehension Test Scores," Journal
of Reading 12 (December 1969): 197.

identified cloze reading levels and cloze criteria as indepen-

dent: 50 percent; instructional: 30 percent; frustration: 20

percent or less.[1] When Alexander correlated scores cloze tests

and informal reading inventories prepared on graded reader

materials, he reported higher criteria levels--independent: 62

percent; instructional: 47 percent; frustration: less than 47

percent.[2]

In another study, in which three measures were compared,

namely scores on an informal reading inventory, the Comprehen-

sive Test of Basic Skills and on a cloze test, Jones and Pikulski

reported the following criteria:[3] independent level could range

from 36 to 50 percent; instructional level from 29 to 44 percent;

frustration level from 17 to 31 percent. They concluded that

the cloze test gave a more accurate reading level placement than

did the standardized test, and that cloze approximated levels on

the informal reading inventory and in its favor required less

time to administer.

Given the findings of Bormuth, Rankin and Culhane, Ransom,

Alexander and Jones and Pikulski, it is apparent that cloze scores

---

[1]Peggy Ransom, "Determining Reading Levels of Elementary
Children by Cloze Testing," p. 481.

[2]H. W. Alexander, "An Investigation of the Cloze Procedure
as a Measuring Device Designed to Identify the Independent, In-
structional and Frustration Reading Levels of Pupils in the
Intermediate Grades," (Doctoral dissertation, University of
Illinois, 1968), cited by Margaret B. Jones and Edna C. Pikulski,
"Cloze for the Classroom," Journal of Reading 17 (March 1974):
435.

[3]Margaret B. Jones and Edna C. Pikulski, "Cloze for the
Classroom," Journal of Reading 17 (March 1974): 437.

are accurate and valid means of establishing independent, instructional and frustration reading levels.

Cloze has also been shown to correlate with other measures of pupil performance such as intelligence, aptitude, mathematical reasoning and verbal skills.

Correlations of Cloze to Intelligence and Aptitude

A number of studies that have been carried out have shown the relationship of cloze to verbal abilities or aptitudes. One such study was Taylor's which correlated a pre-cloze, any-word deletion test with the Air Force Qualification Test for which he obtained a correlation of .72. Other correlations included word knowledge .85 and arithmetical reasoning .70.[1]

Intelligence and a pre-cloze, any-word deletion test were correlated in Jenkinson's study; a correlation of .69 was obtained.[2] Fletcher correlated a pre-cloze test with the American Council of Education Psychological Examination and found a correlation of .72 with the Linguistic subtest and .45 with the Quantitative subtest.[3]

---

[1]W. T. Taylor, "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," p. 24.

[2]Marion D. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension," cited by E. F. Rankin, "The Cloze Procedure--A Survey of Research," p. 139.

[3]J. E. Fletcher, "A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities," cited by E. F. Rankin, "The Cloze Procedure--A Survey of Research," p. 139.

Hafner correlated three measures of intelligence with his pre-cloze scores: Otis Quick Scoring Mental Ability Test yielded a correlation of .73; the Information subtest of the Weschler-Bellevue Intelligence scale yielded a .56; Hafner's General Information Test yielded .46.[1] In addition, the Michigan Vocabulary Profile Test obtained a correlation of .56.

The verbal reasoning tests, Understanding Communication by Thurstone, correlated with two different types of cloze tests and resulted in correlations of .52 and .61 for Greene.[2]

On the negative side, Weaver and Kingston found only a moderate relationship between cloze and verbal comprehension when they submitted eight cloze tests, four written and four oral, and ten cognitive tests to factor analysis.[3] Three factors were identified: verbal comprehension, redundancy utilization and rate memory. Cloze tests were most closely related to redundancy utilization and only moderately related to verbal comprehension.

When Schneyer[4] correlated cloze test scores based on an every-tenth word deletion basis with The California Test of

---

[1]Lawrence E. Hafner, "Relationships of Various Measures to the Cloze," p. 138.

[2]F. P. Greene, "A Modified Cloze Procedure for Assessing Adult Reading Comprehension," (Doctoral dissertation, University of Michigan, 1964), cited by E.F. Rankin, "The Cloze Procedure-- A Survey of Research," p. 140.

[3]W. W. Weaver and A. J. Kingston, "A Factor Analysis of the Cloze Procedure and Other Measures of Reading and Language Ability," p. 259.

[4]Wesley J. Schneyer, "Use of the Cloze Procedure for Improving Reading Comprehension," p. 177.

Mental Maturity and cloze tests based on noun-verb deletions and the same intelligence test, he obtained higher correlations, .63 with the every-tenth word deletion system than with the noun-verb deletion system which yielded .42. This seems to substantiate Rankin's finding that every-nth deletion systems are more highly related to intelligence than are the noun-verb selective deletion systems.

Similar results were obtained with college pupils when Guice correlated cloze scores based on a selective deletion system with intelligence scores.[1] He found only a moderate relationship with a correlation of .40. Cloze correlated much more highly with creativity as tested by the Guilford, Merrifield and Christensen Test of Creativity, "Alternate Uses."

Therefore, depending on the deletion system used, cloze scores correlate quite highly with verbal ability scores. Particularly the any-word deletion system correlates highly to intelligence as opposed to the selective noun-verb deletion system.

It could be concluded that cloze has many possible applications and uses, however, the full potential of the cloze technique has not yet been demonstrated. Few studies have attempted to use cloze as a diagnostic tool for analyzing the reasoning processes of readers.

_____

[1] Billy M. Guice, "The Use of the Cloze Procedure for Improving Reading Comprehension of College Students," Journal of Reading Behavior (Summer 1969): 90.

Cloze as a Diagnostic Tool

The cloze procedure by its very nature would seem to have great potential for determining the differences between skilled and less skilled readers in terms of identifying the strategies and cues used by each. When applying the cloze technique, the reader is engaged in a reasoning or searching process in order to successfully complete the cloze deletions. Ordinarily, this type of activity is performed in reading when the reader is determining a word meaning or unlocking an unknown word. For support, the psycholinguistic theories of reading have advanced the idea that reading may be described in terms of such searching and matching processes and that good readers sample the graphic cues on the basis of their knowledge of the syntactic characteristics of the language.[1] From such an educated guess, the reader checks for semantic appropriateness in its particular context and only if the guess is found to be inappropriate does the reader go back and sample for more cues. The psycholinguists also suggest that good readers use their knowledge of language to a greater advantage than do the inefficient readers.[2]

Hafner's findings tend to document the statements made by the psycholinguists regarding the kinds of processing and

---

[1]Delores K. Kennedy and P. Weener, "Visual and Auditory Training With the Cloze Procedure to Improve Reading and Listening Comprehension," Reading Research Quarterly 8 (Summer 1973): 528.

[2]Ellen B. Ryan and Melvyn I. Semmel, "Reading as a Constructive Language Process," Reading Research Quarterly 5 (Fall 1969): 81.

thinking good readers are engaged in when reading.[1] In his

study in which he investigated the relationship of cloze scores

and various measures such as vocabulary, intelligence, informa-

tion, achievement, personality and reasoning, as well as such

cloze-related variables as cloze speed, incongruency scores and

connective/content ratios, Hafner found that reasoning played

an important role in the execution of cloze tasks. He obtained

positive correlations between cloze scores and the percentage

of grammatically correct responses in cloze responses otherwise

incorrect. Also positive correlations were obtained between

cloze scores and the percentage of correct connective word com-

pletions as compared to content word completions.

Hafner further made the diagnostic observations that

"poor readers" did not use widely separated sections of context

that would aid them in completing cloze items correctly; neither

did they use immediately adjacent context clues.[2] Poor readers

obtained high incongruency scores which were based on the number

of times responses were inappropriate in terms of clearly stated

information in the context, or inappropriate in terms of the lan-

guage patterns of the sentence. Poor readers also had a low

percentage of words that were grammatically correct in responses

that were not exact replacements of the words deleted. In sum-

mary, Hafner concluded that poor readers were: less rapid in

cloze tasks; less able to use contextual clues; less intelligent;

---

[1]Lawrence E. Hafner, "Relationships of Various Measures
to the Cloze," p. 138.

[2]Ibid., p. 141.

more anxious; less knowledgeable; less able to reason well than better readers.

In contrast, those subjects who had higher scores on measures of intelligence, achievement, vocabulary, general reading achievement, information and college academic achievement, completed cloze tasks more rapidly, made fewer incongruent responses and made fewer grammatically incorrect responses in their incorrect responses.[1] The more intelligent and better readers also had a higher percentage of correct connective word completions as compared to content word completions.

In another study, this time with high school pupils, high cloze scorers were found to have superior reading skills. When Jenkinson[2] conducted individual interviews with both low and high cloze scorers and asked subjects to state their reasons for inserting a particular word in a particular blank, she found that the verbalizations distinguished the high scoring subjects as being superior in many important reading skills.

Jenkinson's study indicated that more specific analysis of the actual processes good readers use as opposed to poor readers would be useful information. For example, it would be interesting to know if high cloze scorers use a wider range of clues or more advanced clues. In an attempt to investigate

---

[1]Ibid., p. 140.

[2]Marion E. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension," (Doctoral dissertation, University of Chicago, 1957), cited by Lawrence E. Hafner, "Cloze Procedure," Journal of Reading (May 1966): 419.

this area, Rankin and Overholser looked more specifically at the types of contextual clues that were used by different reading ability groups within the fourth, fifth and sixth grades.[1] These investigators devised a Context Test using the cloze procedure, deleting nouns, adjectives, adverbs and verbs from contexts fulfilling the requirements of each category established by Ames in his classification scheme. Each clue category was illustrated by ten contextual situations with selections for the Context Test taken from two current fourth grade readers. Their findings revealed that the rank order of clue difficulty was very similar for each grade and for each reading level in each grade. The rank order of clues ranging from easiest to most difficult was: words connected in a series, modifying phrases or clauses, language experiences or familiar expression, cause and effect, association, referral clue, preposition, synonym clues definition or description, question and answer, comparison or contrast, main idea or supporting detail, non-restrictive or appositive phrases. The average percentage of accuracy for the Context Test was 43 percent for grade four, 50 percent for grade five, and 62 percent for grade six. The researchers concluded after detailed analysis of the data that whatever differences existed among grade four, five and six with respect to the utilization of context clues could be accounted for by differences in reading ability. They found grade

---

[1]Earl F. Rankin and Betty M. Overholser, "Reaction of Intermediate Grade Children to Contextual Clues," Journal of Reading Behavior 1 (Summer 1969): 59, 61, 71.

placement over and above reading ability an insignificant contribution to the variance associated with each clue. Their investigation showed that intermediate grade children can respond adequately to certain types of clues while finding other types of clues very difficult.

Rankin and Overholser predetermined the types of clues that could be considered by deliberately setting up contextual situations that thirteen of the fourteen different clue categories established by Ames[1] in his study with twenty advanced graduate students. These investigators did not question the subjects to determine whether the subject actually used that particular type of clue to arrive at the answer but merely assumed that was the case if the response was correct. The deletions in the Context Test were limited to content or lexical type words, namely nouns, adverbs and adjectives with no structural or connective words deleted.

Justification for the Study

It has been shown through a review of the literature and research that there is a need to design a study which will overtly question subjects about their reasoning and their use of clues, grammatical, syntactical and semantical, when engaged in completing cloze deletions. Studies reported to date have assumed the use of specific grammatical, language and other contextual clues by virtue of correctly completed cloze items, but

---

[1]Wilbur S. Ames, "The Development of a Classification Scheme of Contextual Aids," Reading Research Quarterly 2 (Fall 1966): 66-7.

have not required actual verbalizations from subjects. From such verbalizations it could be determined whether skilled and unskilled readers are significantly different in their rationale-giving abilities. Also, it could be determined if Hafner's findings with college pupils would be confirmed with grade threes and fives. Would skilled readers at these grades be superior in their ability to complete cloze deletions? Would the subjects who obtain higher cloze scores more frequently provide acceptable rationale responses for cloze items than would subjects with low cloze scores? Would the more intelligent subjects more frequently complete cloze items with exact word replacements? Would boys be better than girls at completing cloze items? If free to initiate reasons and clues, would it be found that subjects would use significantly different clues than those identified in Ames's study? Since cloze measures something different than the context utilization test does, would the clues used by subjects in this study differ significantly from those that were assumed to be used in Rankin and Overholser's study? Could subjects provide rationale for all types of word deletions equally well? These and other questions might be answered if subjects at two elementary grade levels were given cloze tests and then individually interviewed to elicit rationale for their cloze item choices.

CHAPTER 3

DESIGN OF STUDY

The two foci for study in this investigation were: to explore at the third and fifth grades the relationship that exists between the ability to complete cloze items with exact word replacements and the ability to give acceptable rationale for cloze responses; and to explore the relationship that exists between total pre-cloze scores based on exact word replacements and reading ability and intelligence and sex at these two levels. In addition, an item analysis was done to determine the types of items for which acceptable rationale was frequently supplied and the types of items for which the various categories of unacceptable rationale were supplied. As well, an analysis of the total populations at each grade was done to determine which factors identified in this study such as vocabulary, comprehension and intelligence correlated most highly with total cloze scores and with the interview populations at each grade which factors correlated most highly with total rationale scores.

This chapter describes first the pilot study which evaluated the types of cloze tests, testing procedures and the interview format developed for this research study. The research study is then outlined in detail, including a description of the population, the test instruments, the administration of the instruments and the scoring of the instruments. An outline of the

statistical procedures to be used concludes the chapter.

The Pilot Study

Purpose

A pilot study was conducted to evaluate the types of cloze tests, procedures and interview format which would be used in this research study.

Subjects

Nineteen subjects from one suburban school division were chosen as subjects for the first phase of the pilot study. These students were enrolled in grades two to six, ranging in age from seven to eleven. They were selected by the classroom teachers as representing various reading levels.

The second phase of the pilot study involved a total of eight students, four in grade three and four in grade five from the same school. Two subjects from each grade were viewed as low average pupils and two were viewed as high average pupils according to the school record files.

Cloze Tests and Administration

For the first phase of the pilot study, a cloze test was constructed on a selection titled, "The Day I Ran Away," taken from the Reader's Digest Skillbuilder, Level IB, published by Reader's Digest Services Inc., 1965. This particular story was selected because it was short enough so that the story could be used in its entirety. The passage was one hundred and sixty words in length; the Spache readability formula indicated a 1.7 level of difficulty. The first paragraph was left intact

and words were deleted following no particular deletion pattern; instead deletions were dictated by the purpose of having a representative sample of words from the various grammatical classes. There was a total of twenty-six deletions.

Various procedures were tried for administering this test. The first two subjects who were in grade two were taken individually and the investigator read the selection orally. The subjects were then asked to supply the missing words orally while the investigator printed the answers in the blanks. The next two subjects, one each in grade three and four, were taken individually and asked to read the selection orally and supply the answers orally while the investigator wrote the answers. The fifth subject, a grade six student, was taken individually and requested to write in his responses. He read the selection silently first, reread it orally, and printed his answers in the blanks.

Immediately after completing the cloze test, each of the five subjects was interviewed individually and asked to give his/her reason for selecting the words he/she chose for each of the twenty-six cloze items in this test. Responses the subjects gave were recorded by the investigator as they were spoken. A preliminary analysis showed that the pupils could function with such a task in an individual setting and were able to give some kind of rationale for the words they selected. It was essential then to confirm whether the task could be performed as adequately in small groups.

Then fourteen subjects, three in grade five, four in

grade four, two in grade three and five in grade two were taken
in groups of three or four and asked to read the selection si-
lently and to print their responses independently. Spelling
assistance was given if requested. Subjects whispered their
requests and the investigator gave the correct spelling on paper.

Tests were all scored on an exact word replacement basis
and scores on this common cloze test were compared at the various
ages, grades and reading levels. Some limited item analysis was
done to evaluate which types of deletions were found easier by
subjects and which were found most difficult to complete. The
individual interviews were surveyed to decide whether it was
realistic to expect subjects at these primary and early elemen-
tary levels to supply reasons for their cloze responses.

On the basis of this first phase, it was concluded that
the best procedure for completing cloze items was to have pupils
read the cloze test silently and to write their answers indepen-
dently without spelling assistance. It was further concluded
that an every-nth deletion pattern would facilitate the selec-
tion of cloze items and would yield a fair sample of words from
the various grammatical classes. From the responses of these
pilot interviews, it was concluded that subjects at these early
grade levels were quite capable of providing rationale for their
cloze responses, and that the interview format needed to be
standardized to maintain constant conditions for all subjects.
Taping of the individual interviews was seen as a way of per-
mitting more detailed analysis of each subject's response.

Second Phase of Pilot Study

In the second phase, several kinds of cloze tests were developed and administered to four grade three and four grade five students. Each subject was first asked to dictate a story on a topic of his/her own choice. The story was recorded and transcribed and then made into a cloze test by deleting every seventh word, with the first and last sentences being left intact. The readability of each subject's story was determined using the Spache formula for the grade three stories and the Dale-Chall formula for the grade five stories.

A second type of cloze test was developed for each grade based on a selection from a basal reader. The grade three cloze test was constructed on a story titled, "Weighing an Elephant," from the basal reader, Better than Gold, Level 3 in the Harris Clark Series, published by MacMillan, 1964. The first sentence of the story was left intact and an every-seventh, any-word deletion pattern was applied until a total of twenty-six deletions was also left intact. The readability of this passage was estimated to be at a 2.4 level of difficulty using the Spache readability formula.

The grade five cloze test was constructed on a story titled, "Come Dance With Me," from the reader Sounds of Mystery, in the Sounds of Language Series, edited by Bill Martin, Jr., published by Holt, Rinehart and Winston, 1967. The first and last sentences were left intact and an every-seventh, any-word deletion pattern was applied until a total of twenty-nine deletions was made. The readability of this story was estimated

to be at a 5.8 level of difficulty using the Dale-Chall readability formula.

Each subject at the grade three level and at the grade five level was given three different cloze tests. These cloze tests were given in a group setting on three different days. Each subject completed three cloze tests: one based on his or her own story; on another subject's story, and on the basal reader story. Comparisons of exact word replacement scores were made and recorded in tables 3.1 and 3.2.

TABLE 3.1

COMPARISON OF GRADE THREE SUBJECTS ON
THREE DIFFERENT CLOZE TESTS ON THE
BASIS OF EXACT WORD SCORES

Subject Group	Readability Own Story	Own Story	Other's Story	Basal Story (Readability 2.4)
S Low	2.0	81%	15% (G's)	46%
J Low	2.2	95%	64% (S's)	42%
H High	2.5	71%	65% (J's)	62%
G High	3.0	100%	78% (H's)	57%

TABLE 3.2

COMPARISON OF GRADE FIVE SUBJECTS ON
THREE DIFFERENT CLOZE TESTS ON THE
BASIS OF EXACT WORD SCORES

Subject Group	Readability Own Story	Own Story	Other's Story	Basal Story (Readability 5.8)
D Low	5.0	93%	57% (K's)	62%
K Low	4.1	93%	63% (D's)	51%
R High	4.6	85%	62% (M's)	79%
M High	5.2	100%	58% (R's)	69%

As the tables indicate, all subjects scored highest on their own stories when written as cloze tests. Scores on these tests were very high for all subjects with little range between scores. All the grade three subjects, except S, scored second highest on another subject's story written as a cloze test. Only one grade five scored second highest on another subject's story. The scores on the basal story seemed to match the reading levels of the subjects and separated the poorer readers from the better readers.

A second passage with a readability much closer to the grade three subject's grade placement was then selected and written as a cloze test. This cloze test was constructed from a story titled, "Jeffrey and the Airport Gate," from the reader, Stories of Fun and Adventure, published by Copp Clark Publishing Co., 1964. The readability of this story was estimated to be at a 3.7 level of difficulty using the Spache readability formula. An every-seventh, any-word deletion pattern was applied until twenty-eight deletions were made. These scores were compared to the scores obtained on the first basal passage which had a readability of 2.4. Also, an alternate scoring system was tried whereby responses which were syntactically or semantically acceptable were counted. The results are shown in tables 3.3 and 3.4.

The tables indicate that this alternate scoring system, which permitted synonyms or syntactically acceptable responses, resulted in increased percentage scores ranging from seven percent to twenty-one percent. The high cloze scorers in both

grades made the largest gains as a result of this alternate
scoring system.

TABLE 3.3

EXACT WORD SCORES COMPARED TO ACCEPTABLE
WORD SCORES FOR GRADE THREE
(SECOND PASSAGE)

Subject	Exact Word % Readability 3.7	Acceptable Word % Readability 3.7	Increase
S	25%	32%	7%
J	36%	46%	10%
H	28%	50%	22%
G	64%	85%	21%

TABLE 3.4

EXACT WORD SCORES COMPARED TO ACCEPTABLE
WORD SCORES FOR GRADE FIVE
(FIRST PASSAGE)

Subject	Exact Word Readability 5.8	Acceptable Word % Readability 5.8	Increase
D	62%	69%	7%
K	51%	62%	12%
R	79%	96%	17%
M	69%	86%	17%

All eight subjects were interviewed individually on two
separate occasions to elicit the subject's rationale for their
cloze responses made on both the "other subject's story" written
as a cloze test, and on the basal reader cloze test. Subjects
were asked, "Why did you choose that word for that blank?", or,
"What clues did you use to help you pick this word?", or, "How

did you know that this was the best word for that blank?".

Responses were taped and later transcribed for analysis.

Analysis of responses made by the grade threes indic-
ated that they were able to use a variety of clues and reasons
for their choices of words. Responses were categorized into
such headings as use of main ideas or title, reference to pre-
viously stated detail, use of own experiential background, use
of grammatical knowledge involving possessive pronouns, pro-
nouns and past tense and familiarity with figures of speech and
use of dialogue patterns.

Analysis of the grade five responses indicated that
they were able to use a greater variety of clues and reasons
than the pupils in the third grade. Thus the responses were
categorized into such groupings as use of main idea or title,
story content, interpretation of story content, use of question
and answer structure, use of grammatical knowledge including
possessive pronouns, pronouns, noun determiners, conjunctions,
past tenses, use of familiar expressions and use of word order
or syntax.

Evaluation of the Pilot Study

Following the completion of the pilot study, it was
decided that the most feasible type of material upon which to
construct a cloze test was the basal reader material. Sub-
ject's own story material written as a cloze test tended to
yield very high cloze scores providing little range in scores
and not sufficiently discriminating to separate the good

readers from the poor readers. Further, the length of the stories dictated varied from subject to subject with some stories being too short to be realistically used for a cloze test. Generally, the content tended to be of a very personal nature and based on a specific interest of that subject, thereby handicapping another reader's chances of making a valid score. For example, one subject wrote about a hockey experience stating the number of goals scored, the name of the team, name of the tournaments which in turn developed into cloze items. It was found that a passage with a readability at or very near the grade placement level of the subjects yielded the best range in scores and identified the poor readers and good readers at each grade level.

Deleting words on a selective basis did not produce any more information than when words were selected on an every-nth, any-word deletion pattern. It was decided that an every-fifth, any-word deletion pattern would permit a shorter passage to be used than would an every-seventh, any-word selection pattern. Thus the basis of a minimum of fifty items was selected, which would yield a good sample of all classes of words and provide a reasonable number of items for statistical analysis. Exact word scoring was decided on as the most objective and most valid means of scoring.

It was also decided that independent completion of cloze items without spelling assistance would produce the most reliable and uncontaminated results. In addition, it was found that a test read orally by the examiner could provide unintentional

clues through voice or intonation; thus a decision was made to have the subjects read silently and independently.

The interview format tried in the pilot study was standardized to maintain as constant conditions as possible from one interview to the next. It was decided that the best procedure would be to have subjects read silently until the first cloze item and then the investigator would ask the standard question, "Why did you choose this word for this blank?" with the investigator pointing to the cloze item in question. If no response was forthcoming, a second question would follow, "Why did you pick this word?". If no response was given after approximately sixty seconds, the investigator would proceed to the next completed item and repeat the procedure. Rationale would be requested for all items, including those not completed with exact word replacements so that the continuity of the story would not be lost between items. It was also decided that the responses should be taped and later transcribed to permit thorough analysis.

Design and Procedure of Research Study

Description of the Population

Subjects for this study were all the students in grade three and grade five at another school in the same suburban school district in which the pilot study had been conducted. The entire population was chosen from one school to lessen the possibility of a skewed distribution of reading achievement and intelligence. This school is situated in an average socio-economic area and was seen as representing an average population

of pupils at these two grade levels.

There was a total population of 195 subjects with 91 pupils in grade three and 104 pupils in grade five. In grade three, there were 53 boys and 38 girls, all of which remained for the entire research period. At the fifth grade, 5 subjects were dropped because of absenteeism, which left a total of 99 subjects in the study. Out of this total, 45 were boys and 54 were girls. The age range of the sample was from 8 years, 7 months to 13 years, 1 month.

Description of the Test Instruments

Cloze Tests

Two cloze tests were constructed for use in the research study; one to be used with the grade three subjects and one to be used with the grade five subjects. Each test was constructed following the criteria established by previous researchers such as Taylor, Rankin, Bormuth and Ruddell, as well as those criteria selected on the basis of the pilot study.

A passage was selected from a basal reader which had a readability that fell at or very near the grade placement level of the subjects. The passage was at least 250 words in length with exactly 50 cloze deletions. An every-fifth, any-word deletion pattern was applied and standard blanks, 12 typed dashes in length were substituted for the words deleted. First and last sentences were left intact.

The grade three cloze test was constructed from a selection titled, "How Che-Ka-Pash Made the Sun Stand Still," taken from Stories of Fun and Adventure, published by Copp Clark in

1964. The Spache Readability formula was applied and the readability was found to be 3.7, which was very near the subject's grade placement level of 3.9 at that time.

The cloze test for the subjects in grade five was constructed on a selection titled, "History of the Ice Cream Cone," taken from the reader Sounds of Mystery, edited by Bill Martin, Jr., published by Holt Rinehart and Winston in 1967. The Dale-Chall readability formula was applied and the level of difficulty was found to be 5.9, which was at the exact grade placement level for these subjects at that time.

Gates-MacGinitie Reading Tests

The vocabulary and comprehension subtests of the Gates-MacGinitie Reading Tests were administered to secure reading achievement scores for subjects in these two areas. Form I was used in each case with Primary C being administered to grade three, Survey D to grade five. The speed and accuracy subtest was omitted from Survey D.

The Gates-MacGinitie Reading Tests are a revision of the Gates Reading Survey Tests and it is reported that the new edition reflects a marked improvement over its predecessor. There is particularly an improvement in terms of interest correlations with the correlations for the new edition falling substantially below alternate-form reliabilities while in the earlier edition the interest correlations nearly approached the alternate-form reliabilities. The most severe limitation of the earlier edition which was the fact that one instrument was used to measure a wide range of grades, has been eliminated

by the development of several instruments to test this same range. This series of tests was seen as appropriate for this research study because as reported, interest correlations are low enough to permit a rough estimate of skills in two areas of reading, namely vocabulary and comprehension, and are best used as survey tests.

The possible choice of test instruments to match particular grade levels was seen as another advantage in using this series and the conversion of raw scores into grade scores rather than percentiles facilitated the analysis of data.[1]

Otis-Lennon Mental Ability Tests

The Otis-Lennon Mental Ability Tests were administered to obtain measures of intelligence for all subjects. In each case, Form J was used with Elementary I level being administered to grade three, Elementary II level to grade five.

The Otis-Lennon Mental Ability Tests were used because they are reported to be outstanding tests of their kind. The construction and norming of the tests testifies adherence to the highest level of current standards. Substantial evidence is provided to indicate that the Otis-Lennon is highly reliable and care was taken to reduce the effects of reading handicaps on the student's performance with no reading required at the three lower levels and the median reading grade-level ratings for the other three below the lowest grade for which the test was intended.

---

[1]Oscar K. Buros, ed., The Seventh Mental Measurements Yearbook, (Highland Park, New Jersey: The Gryphon Press, 1972), pp. 1080-83.

It is viewed as a good predictor of scholastic success.[1]

Administration of Instruments

All tests were administered by the investigator in the month of June. Testing began with grade three. A timetable was set up scheduling the four grade three classes at different time periods for the administration of the cloze test.[2] A set of directions including a sample paragraph was attached to each cloze test. These directions were read orally by the investigator as the pupils followed along with the printed form. The samples were done orally with the class participating. Questions were answered if any doubts remained as to how to proceed. The cloze test was then done silently and independently without spelling assistance. No time limit was imposed, but it was found that most pupils completed the cloze test in approximately thirty minutes excluding the time taken for the initial instructions.[3]

Following the administration of the cloze test, all items were scored on an exact word replacement basis. The total cloze scores were then ranked from lowest to highest and using a random number table, fifteen of the lowest cloze scorers and fifteen of the highest cloze scorers were slected. These thirty subjects were then interviewed individually for the purpose of requesting each subject to give their reasons for selecting his

---

[1] Ibid., pp. 689-93.

[2] See Appendix B for the timetable schedule.

[3] See Appendix A for the cloze test and directions given to grade three.

or her answers for each cloze item completed. Omitted items were not questioned. The subjects were asked to read silently up to the first cloze item and then the standard question was asked of each subject, "Why did you choose this word for this blank?" with the investigator pointing to the item in question. If no response resulted, a further question was used, "Why did you pick that word?". If after approximately sixty seconds no response was forthcoming, the investigator went on to the next completed item. Subjects read silently between items. Once the process was underway, the question was reduced to, "Why did you choose this word?". Pupil responses were taped and transcribed for analysis purposes.

Once the individual interviews were completed, the Gates-MacGinitie Reading Tests Primary C were administered, which included a vocabulary and reading comprehension subtest. Again, a timetable was set up and each grade three class was scheduled at a specific time.[1] The test was administered according to the directions in the test manual.

The four grade five classes were next scheduled for the administration of the cloze test.[2] Three of the grade five classes were from an open area and the fourth class was from a closed classroom. This closed classroom was used for all the testing sessions. Each of the four cloze testing sessions was conducted in the same manner. For each session, the same set

---

[1]See Appendix B for timetable schedule.

[2]See Appendix B for timetable schedule.

of directions and samples were read orally by the examiner and
the pupils followed along with the written directions. Samples
of cloze items were done with the class orally to ensure that
all subjects understood the nature of the cloze procedure.[1] No
time limit was imposed, however, most subjects completed the
test in approximately thirty minutes excluding the time taken
for the initial directions and samples.

As was done with grade three, the total cloze scores of
the subjects in grade five were ranked from lowest to highest.
Using a random numbers table, fifteen of the lowest cloze sco-
rers and fifteen of the highest cloze scorers were selected for
the individual interviews. The same procedure was followed with
each subject being asked to give his or her reasons for each
cloze response completed. The standard question was, "Why did
you choose this word for this blank?". If a response resulted,
no further questions were asked, however, if no response was
made by the subject a second question was used to stimulate a
response, "Why did you pick that word?". If no response resul-
ted after approximately sixty seconds, the investigator moved
on to the next completed item and the procedure was repeated.
Pupil responses were taped and transcribed for analysis.

When the individual interviews with the grade fives
were completed, a schedule was set up for the administration
of the Gates-MacGinitie Reading Tests, Survey D. The vocabu-
lary and reading comprehension subtests were administered

---

[1]See Appendix A for cloze test and directions given to
grade five.

according to the directions in the test manual. The speed and accuracy subtest was omitted.

Lastly, a schedule was set up to administer the group intelligence tests.[1] First the four classes of grade threes were given the Otis-Lennon Mental Ability Test, Elementary I, Form J. The same procedure was followed as with the reading achievement tests in terms of test setting. The classrooms of each grade three group were used for testing and the one closed grade five room was used for each of the grade five groups. The grade five classes were given the Otis-Lennon Mental Ability Test, Elementary II, Form J. The tests were administered according to the directions in the test manual.

Scoring the Instruments

All the scoring was completed by the investigator. The cloze tests were scored for exact word replacements immediately after being administered and subjects were rank ordered from lowest to highest in terms of total cloze scores. Using a random numbers table, fifteen of the lowest cloze scorers and fifteen of the highest cloze scorers were selected for the individual interviews. Later these cloze scores were rechecked to insure accuracy of total cloze scores. The Gates-MacGinitie vocabulary and comprehension subtests were scored according to directions in the examiner's manual. A count of items correct was checked against a count of items missed to insure correct scoring. Raw scores were converted to grade equivalent scores

---

[1] See Appendix B for the timetable schedule.

and these grade equivalent scores were used in the analysis. Similarly, the Otis Lennon Mental Ability Tests were scored according to the directions in the examiner's manual. A count of items correct was checked against a count of items missed to insure accuracy.

Finally, the rationale responses which had been taped were transcribed and scored. Only the rationale given for those items for which exact word replacements had been made were used for this analysis. In other words, the rationale given for the incorrectly completed items were not examined in this study.

Initially, the rationale responses were graded as either ACCEPTABLE (A), NOT ACCEPTABLE (N), or OMITTED (O). After further and more detailed analysis, it became apparent that two other categories of responses were needed to accommodate two quite frequent types of responses which could not outrightly be graded as acceptable and yet had some merit in terms of rationale. These two categories were the repetition type responses (R1) and the "only one that fits" type response (R2) which both seemed to suggest some intuitive understanding on the part of the subjects. Subjects made these types of responses when they were not able to explain or justify their word choice in a direct statement but yet they had very strong feelings that the word was right.

In the final analysis, there were five possible categories of responses:

1. ACCEPTABLE RESPONSE (A). Responses were categorized

as acceptable if the subject demonstrated the use of a clue or reasoning strategy involving a grammatical clue such as pronoun agreement, possessive agreement or plural agreement; a syntactical or structural clue dealing with word order or phrase structure; a semantical or content clue pertaining to the meaning of the passage such as the main idea, supporting idea or the sequence of events.

2. NOT ACCEPTABLE RESPONSE (N). Responses were categorized as not acceptable if subjects made statements which demonstrated faulty reasoning or no reasoning at all.

3. REPETITION RESPONSES (R1). Responses were categorized as repetitions if subjects repeated or paraphrased the phrase or sentence containing the cloze item, emphasizing the word placed in the blank as if to imply that word fit well in the context.

4. ONLY ONE THAT FITS RESPONSE (R2). Responses were categorized as R2 type responses if subjects responded with one of the following statements or words to this effect:

"It was the only word that fit."

"You couldn't put any other word in there."

"It sounded better than any other word."

"What else could you put in there?"

"What other word would go there?"

"There isn't another word to fit in there."

"I thought it was the only one that made sense."

5. OMITTED (O). Responses were categorized as omitted if subjects gave no answer when rationale was requested.

In summary, each subject's responses were categorized into one of five possible categories: "acceptable," "not acceptable," "repetitions," "only one that fits" or "omitted." The sum total in all their subtotals was equal to the subject's total cloze score.

Statistical Analyses

The statistical procedure employed to test hypotheses one to ten was a nonparmetric test of significant difference between two groups called the Wilcoxon Mann Whitney Test. To further test hypotheses one and two, an analysis of co-variance was done. A descriptive analysis of the items was included as well as a test of multiple regression to determine which factors correlated most highly with total cloze scores and which factors correlated most highly with acceptable rationale scores. These findings were tabulated and are described in the following chapter.

CHAPTER 4

ANALYSES OF THE DATA

The purpose of this study was to obtain empirical data
to determine the relationship between the ability to complete
cloze items with exact word replacements and the ability to give
acceptable rationale for cloze responses, and to determine the
relationship of total cloze scores to vocabulary, comprehension,
intelligence and sex. A second aim of the study was to deter-
mine which of these factors best predicts total cloze scores and
total acceptable rationale scores.

The data obtained was processed through the Health Sci-
ences Computer System via the terminals in the Faculty of Edu-
cation.

The initial phase of the analyses involved plotting the
categories of each subject's rationale responses on an item
analysis grid. This was followed by the tabulation of total
responses in each of the five possible categories for each sub-
ject in the interview populations and a descriptive analysis of
the strategies and cues demonstrated in their rationale.

In the next phase of the analysis, high cloze scorers
and low cloze scorers were compared in terms of total acceptable
cloze responses. A nonparametric test of significant difference
between two independent groups called the Wilcoxon Mann Whitney
Test was used. Null hypotheses one and two were further tested

using the more powerful statistical technique of analysis of
covariance whereby the total acceptable rationale scores were
analyzed with control of total cloze scores.

The next phase of the analyses dealt with the null
hypotheses three through ten regarding the relationship of
total cloze scores with vocabulary, comprehension and intel-
ligence measures as well as sex. The Wilcoxon Mann Whitney
Test, a nonparametric test of significant difference between
two independent groups, was used.

Next a test of multiple regression was used to deter-
mine which predictor variable: vocabulary, comprehension or
intelligence, best predicts total cloze scores, and which pre-
dictor variable: vocabulary, comprehension, intelligence, total
cloze scores best predict acceptable rationale frequency.

In conclusion, a descriptive analysis of the items was
done using the item analysis grids. Frequency tables were made
to identify the words and the categories of words for which
exact word replacements and acceptable rationale were most fre-
quently obtained.

Rationale Response Categories

The category of each subject's rationale response was
plotted on an item analysis grid. "A" indicates an acceptable
rationale credit, "R1" a repetition rationale, "R2" an "only
one that fits" rationale, "N" an unacceptable rationale, and
"0" an omitted rationale.

Table 4.1 shows the categories of the responses obtained
from the low cloze scorers in grade three, and table 4.2 shows

TABLE 4.1

CATEGORIES OF RATIONALE RESPONSES FOR
LOW CLOZE SCORERS--GRADE THREE

ITEMS		SUBJECTS														
	4	9	12	14	16	17	20	25	34	35	37	38	39	40	47
1. curious															
2. always				N	R1				R2	A	R1		A	R1	A
3. trying															
4. things															
5. fell	R1	R1		A		R1			A		R1			A	
6. trying															
7. the	N									N		O			
8. didn't															
9. another															
10. father's															
11. why								R1							
12. day		R1	R1			O	R2	A	A		O	A	N	O	R2
13. the		R1			N		R1			R1	R1		R1	R1	A
14. us															
15. sun			A	A	A	A	A	A	A	N	O	A	O	A	A
16. how							R1			N			O	R1	R1
17. if											R1		A		
18. the		R1	R1	A		O	R1	R1	O	N	O	N	N	R1	O
19. not									N	A			R1	R1	R1
20. bed							R1		A				A	A	
21. the		N					R1	R1	R2						O
22. watched															
23. set								A				O			
24. day	R1	A		A	A				A	A		O	A		R1
25. as															
26. then								N	N	N	O	R1		R1	
27. he			A		A				A	A	R1	O	R1	R1	R1
28. set															
29. to			R1	A		O	N			N	R1		R1		
30. family				A				A	A	A		O			
31. snare															
32. way	N	N				N					R1	R1			
33. about			R1		R2			R1	N	A	R1	R1	R1		
34. thought				R2	O	O				N		O	A	R1	R1
35. so							N								
36. his	R1	A		A	A	R1	A	A	A	N	A	O	A	A	A
37. strand															
38. longest					A										
39. next								R1				A		R2	R1
40. before												O			
41. the		R1				N	R1	R1	O	R1	R1	O			R1
42. tree				A	A					A	R1	A	A		A
43. set															
44. topmost															
45. came														R1	
46. and			N						O	R2	R1	O	N		A
47. time			R1			R1	R1							R1	R2
48. as															
49. sky														R1	A
50. Che-Ka-Pash															

*Acceptable (A)
Repetitions (R1)
"Only One That Fits"(R2)
Not Acceptable (N)
Omitted (O)

TABLE 4.2

CATEGORIES OF RATIONALE RESPONSES FOR
HIGH CLOZE SCORERS--GRADE THREE

	SUBJECTS														
ITEMS	48	56	58	60	62	65	67	68	70	71	72	74	85	88	90
1. curious															A
2. always		R1	A	O		A	A	A		R2	N	A	A	R1	N
3. trying														A	
4. things															
5. fell		R1				A		A	A				A		
6. trying														N	
7. the	O	N			R1								R1		
8. didn't															
9. another															
10. father's															
11. why	N											A			
12. day	R1	A			A	A	A	A	A		A	A	A	R1	A
13. the			N		R2	A	A	A	N	R2		R2	A	N	
14. us								A							
15. sun	A	A	A	A	A	A		A		A	A	A	A	A	A
16. how	R1		R2	O	R1	R2	A	A	A	R1	R2	R1	R2	R2	
17. if	R1		R2	O	N	N	R1	A	O	O	O	R2	R1		R2
18. the		R1	R2	N	R1	R2	R1	R2	N	R2	R1	R2	R1	N	R1
19. not		A	R2	A	R1	A	A	A	O		A	A	R1	N	A
20. bed	R1	A	A	A	A	A	A	A	O	R1	A	A	A	A	A
21. the		R1			N	O	A	R2		R2	R2	R2	R1		R2
22. watched	R1		O	A	A	A	A	R2				A	A	A	A
23. set				A	A		A						A	A	A
24. day	R1	A	A	A	N	A	A	A	A	R2	O	A	A	A	A
25. as															R2
26. then	R1	A	A		R1	N	A			N	A	R2		O	A
27. he	R1	A	A	A	A	A	A	A		A	A		A	A	A
28. set									A						
29. to	N	A		N		N	N	R1	A	O	A	R2	R1	O	N
30. family		A					A	A	A	A	A	A	A	A	A
31. snare															
32. way				O	A		A	A	A	O			R1	O	N
33. about				O	A		A	A	A	R2	O	R2	R1	N	
34. thought		R1		A	R1		A	A	A	O	A	A	R1	R1	A
35. so															N
36. his	R1	A	A	A	A	A	A	A	A	A	A	A	A	A	A
37. strand															
38. longest			A	A						O	A			A	A
39. next	R1	A			A	R2			A		A		R1	A	N
40. before		A							A						
41. the	R1		R2	N	R1	O		R1	R1	R2	R1	R2	R1	O	N
42. tree	N		A	A		A		A	A	A	A	A		A	A
43. set		A							A						N
44. topmost															
45. came												R2			
46. and	O	A	A	A	A		A			R2	O	A	R1		R2
47. time	O		N		R1				A	R1				A	
48. as															
49. sky			O									A			A
50. Che-Ka-Pash														A	

*Acceptable (A)
 Repetitions (R1)
"Only One That Fits" (R2)
 Not Acceptable (N)
 Omitted (O)

the categories of the responses obtained from the high cloze scorers in grade three.

Table 4.3 shows the results of the low cloze scorers in grade five, and table 4.4 the high cloze scorers in grade five.

Each subject's responses were totalled for each of the five categories. Table 4.5 shows the totals for grade three and Table 4.6 shows the totals for grade five.

This is followed by a descriptive analysis of the rationale responses generated by the subjects.

Descriptive Analysis of
Rationale Responses

Strategies and Cues Used in Rationale Responses

Subjects in the interview populations demonstrated in their rationale responses that pupils in grade three and five used a wide range of cues and employ a number of strategies when responding to cloze items.

One strategy used by subjects in both grades was giving other options in the same category or class or words and then reasoning that the one they had chosen was the most appropriate.

For example, a subject in grade three gave the following rationale for the cloze item number 12, day:

"I thought it sounded better than night or hour or month or all those [words]."

Other examples involving more advanced categories were the response to item number 21, the:

"Cause in that days that followed, it doesn't sound right."

and for item number 2, always:

TABLE 4.3

CATEGORIES OF RATIONALE RESPONSES FOR
LOW CLOZE SCORERS--GRADE FIVE

	ITEMS								SUBJECTS							
		2	5	10	13	19	20	31	33	37	39	41	42	43	44	48
1.	has		O		R2	R2		N	R1	R1	N	R2	R2		N	
2.	of	O		N	R2	R2		R1	R1		N	N	R2		R2	
3.	but		R1		R2	R1	R1	R1			O					
4.	until															
5.	century									R1						
6.	street															
7.	glass								R1							A
8.	great					A				A	A	A	A		A	A
9.	care															
10.	breaking							A			A	A	A			
11.	away										R2		A			
12.	and									R1	O					
13.	ice	A	A		A		A	A	A	A	A	A	A	A	A	A
14.	when						A	A		R1	A	R1		R1		A
15.	cone	R1				A	A	A	A	A	N		A	A		A
16.	was		R1	R1	R2	R1	R1	R1	A	R1		R1	R1	R1	R1	R1
17.	who										R1			R1		N
18.	a	O				R1		R1	R1	A				R1	N	
19.	of		R1	R1		A	A	R1	A	R1	N		R1	R1	A	
20.	Marchiony		A	A	A	A		A	A	A	A	A	A	A	A	A
21.	his															
22.	new	A		A					R1				R1	R1	A	A
23.	that				R2			R1					R1	R1		R1
24.	he	O	A		A		A	A	A			A	A	A	A	
25.	new			A			A			A		A			A	A
26.	ice	A	R1	A	A	A	A		A	A	A	A	A	A	A	A
27.	thin															
28.	up		A				A									
29.	a			R1	R2	R1	R1	R1	R1	A		R1	R1	R1	N	A
30.	ice	A	R1	A	A	A	A		A	A	A	A	A	A	A	A
31.	his															
32.	his															
33.	ice		A	A	A	A	A	A	A	A	A	A	A	A	A	A
34.	was			N	R1	R1	R1	R1	R1	N	R1	R1	R1	R1	R1	N
35.	of							R1		R1	R1					
36.	edible															
37.	popular			A	A							A				
38.	were		R1	N	R2	R1			A				R1	R1	R1	R1
39.	because		R1			R1	A	N		A		R1	A	R1	A	A
40.	Mr.		A	A	A		A	A	A	A	A	A	A	A	A	A
41.	in				R1	A	A		A	A	N	A	A	A		
42.	toot													A	A	A
43.	idea															
44.	never															
45.	rest												A			A
46.	wasn't															
47.	world's															
48.	the						R1	R1			O	R1		R1		R2
49.	introduced															
50.	so											O			N	

*Acceptable (A)
Repetitions (R1)
"Only One That Fits" (R2)
Not Acceptable (N)
Omitted (O)

TABLE 4.4

CATEGORIES OF RATIONALE RESPONSES FOR
HIGH CLOZE SCORERS--GRADE FIVE

	ITEMS	56	57	59	60	62	68	70	78	81	83	84	89	92	97	98
														SUBJECTS		
1.	has		R1	O	R2	A			A	N				N	R1	N
2.	of	N		A	R2	R2	R1	O	R1	R1	R2	R2	O	R1	R1	N
3.	but	R2	A	R1		R2		R1				O		A	R1	N
4.	until															
5.	century				O							A	R1		R1	R1
6.	street												A			
7.	glass			A					A	A	A	A	A	A	A	
8.	great	A	A	A	A	A	A	A		A	A		A	A	A	
9.	care															
10.	breadking				A	A	A	O				A	N	A	A	A
11.	away						A	A	A	A	R2	R2	A	A	A	A
12.	and									N			N	R2		
13.	ice	N	A	A	A	A	A	A	N	A	A	A	A	A	A	A
14.	when	R2	A						R1		A	R1		A	A	A
15.	cone	A	A		A	A	A	A	A	A	A	A	A	R1	A	A
16.	was	R2	R1	R1	A	N	R1	R1	R1	R1	R1	A	O	R1	N	R2
17.	who				R2		R1	R1					O			N
18.	a	A	A	R1		N	A		A		A	A		R1		A
18.	of	R2	R1	A	R2	A		A	A	R1	A		A	A	A	N
20.	Marchiony	A	A	A	A	A	A	A	A	A	A		A	A	A	A
21.	his															
22.	new				A	A		A	A		A	A	A		A	A
23.	that	R1	R1		R2		R1	O			N	R1	O	R1	R1	N
24.	he	A	A		A	A	A		A	A	A	A	A	A	A	A
25.	new					A		A	A	A		A		A	A	
26.	ice	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
27.	thin															
28.	up		A	A	A		A	A			A	A				A
29.	a	R2	A		R2	A		A	A	A	A	A	R2	N	R1	A
30.	ice	N	A	A	A	A	A	A	A	A	A	A	A	A	A	A
31.	his										A					
32.	his		A													
33.	ice	N	A	A	A	A	A	A	A	A	A	A	A	A	A	A
34.	was	R2	R1	R1	R2	A	R1	R1	R1	R1	R1	A	A	R1	R1	A
35.	of			R1	R2		R1		R1	N		R1		R1	R1	R1
36.	edible															
37.	popular			A	A			A	A		A		O	A	A	A
38.	were			R1	R2	A	A	R1	R1	A	R1	R1	O	A	R1	
39.	because		A	A		A	R1	O	A	A	R1	O	A	R1	A	A
40.	Mr.	A	A	A	A	A	A	A	A	A	A		A	A	A	A
41.	in	R2	A	A		A	A	A	A	A	A	A	A	A	A	A
42.	root		A	A		A	A		A	A			A	A	A	
43.	idea															A
44.	never															
45.	rest	A				A								A		
46.	wasn't										R1		A	R1	A	
47.	world's															
48.	the	R2	R1	R1		R1	R1		N	A	N	O	O	A	R1	A
49.	introduced									A				A		
50.	so														R1	R2

*Acceptable (A)
Repetitions (R1)
"Only One That Fits" (R2)
Not Acceptable (N)
Omitted (O)

TABLE 4.5

TOTAL RATIONALE RESPONSES IN EACH
CATEGORY--GRADE THREE

SUBJECTS	ACCEPTABLE (A)	REPETITIONS (R1)	ONLY ONE (R2)	NOT ACCEPTABLE (N)	OMITTED (O)	TOTAL (T)
4	0	3	0	2	0	5
9	2	5	0	2	0	9
12	2	5	0	1	0	8
14	8	0	1	1	0	10
16	6	1	1	1	1	10
17	1	3	0	2	4	10
20	2	6	1	2	0	11
25	5	7	0	1	0	13
34	8	0	2	3	3	16
35	7	2	1	7	0	17
37	1	11	0	1	4	17
38	4	3	0	1	10	18
39	7	5	0	3	2	17
40	4	11	1	0	1	17
47	7	7	2	0	2	18
48	1	11	0	3	3	18
56	14	5	0	1	0	20
58	10	0	5	2	3	20
60	13	0	0	3	4	20
62	10	8	1	3	0	22
65	12	0	3	3	3	21
67	18	2	0	1	0	21
68	17	2	3	0	0	22
70	16	1	0	2	3	22
71	5	2	9	1	5	22
72	12	3	1	2	4	22
74	13	1	9	0	0	23
85	12	12	1	0	0	25
88	14	3	1	5	4	27
90	16	1	4	7	0	28

TABLE 4.6

TOTAL RATIONALE RESPONSES IN EACH
CATEGORY--GRADE FIVE

SUBJECTS	ACCEPTABLE (A)	REPETITIONS (R1)	ONLY ONE (R2)	NOT ACCEPTABLE (N)	OMITTED (O)	TOTAL (T)
2	4	1	0	0	3	8
5	6	7	0	0	1	14
10	8	3	0	3	0	14
13	8	2	6	0	0	16
19	8	7	2	0	0	17
20	13	5	0	0	0	18
31	8	9	0	2	0	19
33	12	8	0	0	0	20
37	13	6	0	1	0	20
39	9	3	1	5	3	21
41	12	6	1	1	1	21
42	14	6	1	0	0	21
43	10	11	0	0	0	21
44	13	3	1	4	0	21
48	16	3	1	2	0	22
56	8	1	9	4	0	22
57	18	6	0	0	0	24
59	15	7	0	0	1	23
60	14	0	9	0	1	24
62	21	1	2	2	0	26
68	16	8	0	0	0	24
70	16	5	0	0	4	25
78	19	6	0	2	0	27
81	19	5	0	3	0	27
83	19	4	2	2	0	27
84	19	4	2	0	3	28
89	18	1	1	2	7	29
92	21	9	1	2	0	33
97	21	11	0	1	0	33
98	22	2	2	6	0	32

"Because <u>sometimes</u> wouldn't sound right."

Grade five subjects used this same strategy but less frequently for example, in response to cloze item number 7, <u>glass</u>:

"Because they were saying they were getting broken and stuff, so <u>plastic</u> doesn't get broken."

Another example is the response to item number 22, <u>new</u>:

"Well,...you wouldn't put buying old ones...if he broke them he wouldn't buy old ones, he'd have to buy new ones."

Another frequently used strategy was the repetition of the sentence or phrase containing the cloze item with strong vocal emphasis on the cloze item followed by a statement to the effect that the word chosen was the only or most appropriate word possible. An example of this was a response for item number 17, <u>if</u>:

"I read a part of this sentence, <u>blank</u> it were daylight all the time, so it would have to be <u>if</u>."

Or subjects would reply:

"It's the only one that makes sense."

These last two strategies were used when subjects seemed to have an intuitive feel for the correctness of their choice but were unable to explain technically why it was correct.

Another strategy was the use of retroactive context where a subject used a later cloze item to help them choose the answer for a preceding cloze item. For example, one grade three subject said in response to item 22, <u>watched</u>:

Because before that it said <u>sun</u>, I said <u>sun</u> and before I got that sun answer I was going to leave that out but then I read this and I got that [answer]."

Also subjects used immediate context surrounding the word and

more distant context to help them in the selection of words for cloze items. An example of distant context was when a subject gave the following response for item 27, <u>he</u>:

> "Che-Ka-Pash is a.......it says back in the story that Che-Ka-Pash was a boy not a she."

Subjects used a variety of reasons when giving rationale for cloze items. Included in the responses were references to grammatical clues such as pronoun agreement, possessive pronoun agreement, past tense, predicate agreement, article or noun determiners, conjunctions, plural forms, abbreviations, prepositions; references to semantical clues using main idea, story detail, story characters, subject's own background of experience, knowledge of expressions, and syntactical clues referring to word order and sense of sentence structure.

Examples of the grammatical clues are:

Pronoun Agreement

Grade 3, Item 27:

> "Well <u>he</u> would fit because it's about a little Indian boy."

Grade 5, Item 24:

> "Cause, they were Mr. Marchiony's customers so.....<u>his</u> customers."

Past Tense

Grade 3, Item 5:

> "Cause it wouldn't make sense if it said once he <u>fall</u> into the river."

Grade 3, Item 34:

> "Cause it wouldn't be good to put the more he <u>think</u> it might be done, doesn't make sense."

Grade 5, Item 16:

"You usually say <u>was</u> when they're supposed to be dead or something, this <u>was</u> in Eighteen Hundred something."

Grade 5, Item 34:

"Because it was from the past tense, so you wouldn't say <u>is</u> glad to be free, you'd say <u>was</u> glad to be free."

Predicate Agreement

Grade 5, Item 16:

"I guess that was the only one that fit, it couldn't be <u>were</u> or something cause there's only one inventor."

Grade 5, Item 38:

'You couldn't put they <u>was</u> called toot cones, it doesn't sound good, not the proper way."

Noun Determiners

Grade 5, Item 18:

"Cause when there's a consonant you put 'a', when there's a vowel you put 'an'."

Grade 5, Item 18:

"Cause it was from <u>a</u> pushcart, not more than one, just one."

Grade 5, Item 18:

"Cause it's a thing you used to put <u>a</u> in front of a word, it's sort of like you say you're pushing <u>a</u> something, or you're taking <u>a</u> walk."

Grade 5, Item 18:

"Well you would say he had pushed it from a certain cart, it was <u>a</u> cart, he wouldn't say some carts."

Grade 5, Item 18:

"Well it was sort of <u>the</u> ice cream, <u>the</u> little shoe, sort of like a title for it,"

Conjunction

Grade 5, Item 12:

"Cause <u>and</u> is a conjunction or whatever and it puts these two together."

Abbreviations

Grade 5, Item 40:

"Well <u>Mr</u>. Marchiony...., well I wouldn't put Mrs. because he's a he."

Prepositions

Grade 5, Item 19:

"Streets of New York, you could put <u>in</u> there too, put <u>of</u> too, cause a pushcart in the streets of New York City."

Examples of semantical clues used are:

Reference to Main Idea

Grade 3, Item 34:

"Well it could have been done but it seems impossible to do it, to catch the sun, he just sat there and <u>thought</u>."

Grade 5, Item 13:

"Cause the whole thing's about <u>ice</u> cream, it just says right here, says cream, so ice cream."

Grade 5, Item 33:

"Well that's probably the main idea of the story."

Reference to Story Detail

Grade 3, Item 22:

"Because he was sitting and watching the sun rise and set, it said before."

Grade 5, Item 15:

"Well up here, it said he baked something in a cone shape, so I put ice cream cone and we have ice cream cones now."

Reference to Story Characters

Grade 3, Item 36:

"Like he asked his sister, wasn't his friends or anything."

Grade 3, Item 27:

"Like he had seen his father make a snare...."

Grade 5, Item 40:

"Well it's telling about Mr. Marchiony and it didn't tell about Miss Marchiony or Mrs. Marchiony, the invention was about Mr. Marchiony."

Use of Subject's Own Background Information

Grade 3, Item 30:

"Well tribe would be a whole lot of families together, family would be better [than tribe] because the other fathers in the tribe would go hunting for their families, there's no reason why Che-Ka-Pash's father should go hunting for the others."

Grade 3, Item 38:

"Well the shortest wouldn't actually be as good as longest because you need a pretty big piece of hair to set a snare for the sun."

Grade 5, Item 37:

"Cause they sort of come...um...they were new and something that's new usually becomes popular."

Grade 3, Item 5:

"He couldn't swim in a cold river because he would freeze to death."

Knowledge of Expressions

Grade 3, Item 12:

"Because well, say when you start a sentence it says one day."

Grade 5, Item 5:

"Cause usually on the radio or something they say at the turn of the century, they wouldn't say at the turn of the day."

Grade 5, Item 19:

"Well if you put streets in New York, it wouldn't really
sound right, but of New York, if you watch shows on T.V.
they usually talk about streets of New York."

Use of Logic and/or Absurd Alternative

Grade 3, Item 42:

"It wouldn't sound right if he climbed to the top of the
tallest bush."

Grade 3, Item 14:

"Cause it was us, they were talking about not, it was March
to us, not them because sometime they believed it was a
different...they didn't have months and that then."

Grade 3, Item 28:

"Well how would father trap snares in the bush, that doesn't
make sense, if he would trap the snares, they're already
traps."

Examples of syntactical clues used:

Juxtaposition of Words

Grade 3, Item 15:

"Because it says there rise, sunrise cause the sun rises up."

Grade 3, Item 15:

"Because the sun doesn't rise and rise, it rises and sets."

Grade 5, Item 26:

"Cause of cream, hardly anything else would go with cream, so
ice cream."

Grade 5, Item 30:

"Ice, it belongs with cream."

Grade 5, Item 23:

"Cause it sounded good in front of 'had been broken'."

Sentence Structure

Grade 3, Item 33:

"Like about--you don't think <u>at</u> it, you think <u>about</u> it."

Grade 3, Item 19:

"It says here then, I would blank have to go to bed, so it would have to be <u>not</u>."

Grade 5, Item 17:

"Cause he sold it, cause <u>who</u> sold ice cream, he's the one <u>who</u> sold it."

Word Function

Grade 5, Item 39:

"Oh <u>because</u> you usually give a reason [with it] like then they have <u>resembled</u>, <u>because</u> they resembled little horns."

Grade 5, Item 39:

"Well it's just the word is explaining why it's called that...<u>because</u>."

Grade 3, Item 26:

"It wouldn't sound right...and <u>the</u> Che-Ka-Pash had an idea...<u>then</u> Che-Ka-Pash had an idea."

This is only a representative list of the strategies and clues used by subjects in grade three and five to illustrate the range and variety of rationale responses. Further discussion of items will be included in the last section of this chapter.

Testing of the Hypotheses

Ten null hypotheses were formulated for testing the two main questions of this investigation. One of the main areas was to determine the relationship between the ability to complete cloze items with exact word replacements and the ability to give acceptable rationale for these cloze responses. The second focus

was on the relationship of pre-cloze scores with a measure of vocabulary, comprehension and intelligence and with sex.

Hypothesis 1

There is no significant difference between low and high cloze scorers at grade three regarding the frequency with which acceptable rationale responses are given for cloze items.

Table 4.7 represents the results of the Wilcoxon Mann Whitney Test which was used to determine whether there was a significant difference between the low and high cloze scorers with regard to acceptable rationale responses. For samples of $n_1 = 15$, $n_2 = 15$, the critical value of U for a one-tailed test is ≤ 56, if it is to be significant at the .01 level.

TABLE 4.7

TEST OF SIGNIFICANT DIFFERENCE OF
ACCEPTABLE RATIONALE SCORES BETWEEN
LOW AND HIGH CLOZE SCORERS--
GRADE THREE

Groups	N	Rank	U	Prob. U	Prob. Z
Low Cloze	15	139.5 (lower)	19.5	.000027*	.000110
High Cloze	15	325.5 (upper)			

*Significant beyond the .01 level.

The value of U obtained for low cloze scorers versus high cloze scorers was 19.5, which was less than the critical value of 56 necessary for significance at the .01 level. On the basis of this finding, the null hypothesis was rejected.

A more powerful statistical technique, the analysis of co-variance, was then used to test hypothesis one. This design

was used to control for the initial difference in the subject's total cloze scores which might compound the differences in total rationale scores between the two groups. Tables 4.8 and 4.9 summarize the findings.

TABLE 4.8

ANALYSIS OF CO-VARIANCE OF ACCEPTABLE
RATIONALE SCORES BETWEEN LOW AND
HIGH CLOZE SCORERS--GRADE THREE

Equality of Slope	Equality of Intercept	Mult. T-Test
F (1, 26)	F (1, 27)	(1:2)
.858	3.686	-1.920*

*Significant at the .05 level.

TABLE 4.9

MEAN TOTAL CLOZE, MEAN AND ADJUSTED MEAN
FOR ACCEPTABLE RATIONALE--GRADE THREE

		Acceptable Rationale	
	Total Cloze M	Obtained M	Adjusted M
Low Cloze	13.067	4.267	6.226
High Cloze	22.200	12.200	10.241

The F ratio for the test of equality of slope is .858 for 1 and 26 degrees of freedom, which is not significant. The F ratio for the test of equality of intercept is 3.686 for 1 and 27 degrees of freedom, which is significant.

The obtained mean in acceptable rationale was adjusted upward for the low cloze scorers from 4.267 to 6.226 while the obtained mean for the high cloze scorers was adjusted downward

from 12.200 to 10.241.

When the multiple T-test was applied to the adjusted means, a significant difference at the .05 level of confidence was found; therefore, hypothesis one was rejected. There was a significant difference in the acceptable rationale scores of low and high cloze scorers at grade three which favored the high cloze scorers.

Hypothesis 2

There is no significant difference between low and high cloze scorers at grade five regarding the frequency with which acceptable rationale responses are given for cloze items.

Table 4.10 represents the results of the Wilcoxon Mann Whitney Test which was used to determine whether there was a significant difference between the low and high cloze scorers' acceptable rationale scores. For samples of n_1 = 15 and n_2 = 15, the critical value of U for a one-tailed test is $\leq$ 56 if it is to be significant at the .01 level.

TABLE 4.10

TEST OF SIGNIFICANT DIFFERENCE OF
ACCEPTABLE RATIONALE SCORES
BETWEEN LOW AND HIGH CLOZE
SCORERS--GRADE FIVE

Groups	N	Rank	U	Prob. U	Prob. Z
Low Cloze	15	134.5	14.5	.000007*	.000044
High Cloze	15	330.5			

*Significant beyond the .01 level.

The value of U obtained for low cloze scorers versus high cloze scorers was 14.5 which was less than the critical value of 56

necessary for significance at the .01 level. On the basis of this finding, the null hypothesis was rejected.

Again, the more powerful statistical technique, namely analysis of co-variance, was used to test further hypothesis 2. Tables 4.11 and 4.12 summarize the findings.

TABLE 4.11

ANALYSIS OF CO-VARIANCE OF ACCEPTABLE
RATIONALE SCORES BETWEEN LOW AND
HIGH CLOZE SCORERS--GRADE FIVE

Equality of Slope	Equality of Intercept	Mult. T-Test
$F\ (1,\ 26)$	$F\ (1,\ 27)$	(1:2)
.140	.636	-0.798 (n.s.)

TABLE 4.12

MEAN TOTAL CLOZE, MEAN AND ADJUSTED MEAN
FOR ACCEPTABLE RATIONALE--GRADE FIVE

	Total Cloze M	Acceptable Rationale	
		Obtained M	Adjusted M
Low Cloze	18.200	10.267	13.507
High Cloze	26.933	17.733	14.493

Hypothesis two was accepted on the basis of the calculated F ratios and the multiple T-test results. Both F ratios for the equality of slope and the equality of intercept were not significant.

The obtained mean in acceptable rationale for low cloze scorers was adjusted upward from 10.267 to 13.507 while the ob-

tained mean for high cloze scores was adjusted downward from 17.733 to 14.483.

When the multiple T-test was conducted on these adjusted means, it was determined that there was not a significant difference between low and high cloze scorers with regard to acceptable rationale scores. Hypothesis two was therefore accepted.

Hypothesis 3

There is no significant difference in the total cloze scores of low and high vocabulary scorers in grade three.

Table 4.13 presents the U and Z values and the probability of Z as related to hypothesis three.

TABLE 4.13

HIGH VERSUS LOW VOCABULARY SCORERS IN
TOTAL CLOZE SCORES--GRADE THREE

Groups	N	Rank	U	Z	Prob. Z
Low Voc.	46	1559	478	-4.4291	.000010*
High Voc.	45	2627			

*Significant beyond the .01 level.

The results were significant beyond the .01 level of confidence; therefore, the null hypothesis was rejected. There was a significant difference in the total cloze scores of low and high vocabulary scorers in grade three which favored the high vocabulary scorers.

Hypothesis 4

There is no significant difference in the total cloze scores of low and high vocabulary scorers in grade five.

Table 4.14 presents the U and Z values and the probability of Z as related to hypothesis four.

TABLE 4.14

HIGH VERSUS LOW VOCABULARY SCORERS IN
TOTAL CLOZE SCORES--GRADE FIVE

Groups	N	Rank	U	Z	Prob. Z
Low Voc.	50	1837.5	562.5	−4.645	.000003*
High Voc.	49	3112.5			

*Significant beyond the .01 level.

The results were significant beyond the .01 level of confidence and therefore, the null hypothesis was rejected. There was a significant difference in the total cloze scores of low and high vocabulary scorers in grade five, which favored the high vocabulary scorers.

Hypothesis 5

There is no significant difference in the total cloze scores of low and high comprehension scorers in grade three.

Table 4.15 presents the U and Z values and the probability of Z, as related to hypothesis five.

TABLE 4.15

HIGH VERSUS LOW COMPREHENSION SCORERS IN
TOTAL CLOZE SCORES--GRADE THREE

Groups	N	Rank	U	Z	Prob. Z
Low Comp.	46	1445.5	364.5	−5.3317	.000000*
High Comp.	45	2740.5			

*Significant beyond the .01 level.

Hypothesis five was rejected since the results of the U test were significant well beyond the .01 level of confidence. There was a significant difference in the total cloze scores of low and high comprehension scorers in grade three, which favored the high comprehension scorers.

Hypothesis 6

There is no significant difference in the total cloze scores of low and high comprehension scorers in grade five.

The values of U, Z and the probability of Z are presented in table 4.16 as related to hypothesis six.

TABLE 4.16

HIGH VERSUS LOW COMPREHENSION SCORERS IN
TOTAL CLOZE SCORES--GRADE FIVE

Groups	N	Rank	U	Z	Prob. Z
Low Comp.	50	1454	179	−7.3349	.000000*
High Comp.	49	3496			

*Significant beyond the .01 level.

The results of the U test show that there was a very low probability that hypothesis six was true; therefore, hypothesis six was rejected. There was a significant difference in the total cloze scores of low and high comprehension scorers in grade five which favored the high comprehension scorers.

Hypothesis 7

There is no significant difference in the total cloze scores of grade threes with high and low intelligence quotients.

Table 4.17 presents the values of U and Z and the prob-

ability of Z, as related to hypothesis seven.

TABLE 4.17

HIGH INTELLIGENCE VERSUS LOW INTELLIGENCE
IN TOTAL CLOZE SCORES--GRADE THREE

Groups	N	Rank	U	Z	Prob. Z
Low Intell.	46	1511.5	430.5	-4.8069	.000001*
High Intell.	45	2674.5			

*Significant beyond the .01 level.

Results were significant beyond the .01 level of confidence; therefore, hypothesis seven was rejected. It was found that there was a significant difference in the total cloze scores of grade threes with high and low intelligence quotients which favored the group with high intelligence quotients.

Hypothesis 8

There is no significant difference in the total cloze scores of grade fives with high and low intelligence quotients.

Hypothesis eight was rejected as results were significant beyond the .01 level, as shown in table 4.18.

TABLE 4.18

HIGH INTELLIGENCE VERSUS LOW INTELLIGENCE
IN TOTAL CLOZE SCORES--GRADE FIVE

Groups	N	Rank	U	Z	Prob. Z
Low Intell.	50	1682.5	407.5	-5.7326	.000000*
High Intell.	49	3267.5			

*Significant beyond the .01 level.

It was found that there was a significant difference

between grade fives with high and low intelligence quotients when it came to completing cloze items with exact word replacements, which favored the group with high intelligence quotients.

Hypothesis 9

There is no significant difference in the total cloze scores of boys and girls in grade three.

As table 4.19 shows, the results were not significant, therefore, hypothesis nine was accepted.

TABLE 4.19

BOYS VERSUS GIRLS IN TOTAL CLOZE
SCORES--GRADE THREE

Groups	N	Rank	U	Z	Prob. Z
Girls	38	1968	787	-1.7	.076138 (n.s.)
Boys	53	2218			

There was no significant difference in the total cloze scores of boys and girls at the grade three level.

Hypothesis 10

There is no significant difference in the total cloze scores of boys and girls in grade five.

The results of the U test are summarized in table 4.20 with calculated U, Z and probability of Z.

TABLE 4.20

BOYS VERSUS GIRLS IN TOTAL CLOZE
SCORES--GRADE FIVE

Groups	N	Rank	U	Z	Prob. Z
Girls	54	2994	921	-2.07	.038443 (n.s.)
Boys	45	1956			

Hypothesis ten was accepted since no significant difference was found between boys and girls in grade five with regard to their ability to complete cloze items with exact word replacements.

Analysis of Rationale Categories

Acceptable rationale scores of low and high cloze scorers were compared in hypotheses one and two. It was found that there was a significant difference in the ability of low and high cloze scorers, at grade three but not at grade five, in terms of providing acceptable rationale responses for cloze items.

Further comparisons were made between low and high cloze scorers at each grade level in relation to the other four types of rationale responses, namely repetition--R1; "only one that fits"--R2; not acceptable and omitted categories. When the Wilcoxon Mann Whitney Test was applied it was found there were no significant differences in either grade, when low and high cloze scorers were compared with regard to the frequency with which they generated the other four types of rationale responses.

Test of Multiple Regression

In order to determine which predictor variable was the most relevant in predicting acceptable rationale scores and total cloze scores, a test of multiple regression was applied. The criterion variables were identified as the subject's acceptable rationale score and total cloze score while the five predictor variables were comprehension scores, vocabulary scores,

intelligence quotients, R1 rationale scores and R2 rationale
scores.

Table 4.21 presents the means and standard deviations
of the variables for the interview population in grade three.

TABLE 4.21

MEANS, STANDARD DEVIATIONS OF VARIABLES
FOR INTERVIEW POPULATION--GRADE THREE

Variable	M	S.D.
1. Acceptable Rationale Scores (A)	8.2	5.4
2. R1 Rationale Scores (R1)	4.0	3.6
3. R2 Rationale Scores (R2)	1.5	2.4
4. Total Cloze Scores (T.C.)	17.6	5.8
5. Vocabulary Score (V)	4.1	.9
6. Comprehension Score (C)	4.0	1.2
7. Intelligence Quotient (I.Q.)	102.6	12.8

The mean for acceptable rationale was 8.2 with a total
cloze score mean of 17.6. The mean grade equivalent score in
vocabulary was 4.1 and 4.0 in comprehension. Mean I.Q. was 102.6.
Repetition rationale scores had a mean of 4 with "the only one
that fit" responses meaning at 1.5.

Table 4.22 presents the correlation coefficients between
the variables identified for the interview population. The
highest correlation coefficient is .76 which is between total
cloze scores and acceptable rationale, with a correlation of .71
between acceptable rationale and comprehension.

In order to compare the regression coefficients of each
of the predictor variables, it was necessary to convert the

regression coefficients into beta weights. The predictor vari-
able with the largest beta weight is the best predictor. Table
4.23 presents the findings from these conversions as well as
the results of the test.

TABLE 4.22

CORRELATION COEFFICIENTS FOR VARIABLES
FOR INTERVIEW POPULATION--GRADE THREE

	A	R1	R2	T.C.	V	C	I.Q.
A	1.00						
R1	-.41	1.00					
R2	.24	-.34	1.00				
T.C.	.76	-.04	.40	1.00			
V	.58	-.04	.22	.63	1.00		
C	.71	-.18	.14	.68	.66	1.00	
I.Q.	.61	-.39	.32	.37	.35	.36	1.00

TABLE 4.23

BETA WEIGHTS OF PREDICTOR VARIABLES--
GRADE THREE

	Beta Weights for Predictors		
Criterion Variable	V	C	I.Q.
Acceptable Rationale	.114	.491	.393
T	.773 (n.s.)	3.144*	3.133**
Criterion Variable	A	R1	R2
Total Cloze Scores	.855	.437	.351
T	8.293**	4.097**	3.516**
Criterion Variable	V	C	I.Q.
Total Cloze Scores	.288	.454	.109
T	1.585 (n.s.)	2.494*	.744 (n.s.)

*Significant at the .05 level.
**Significant at the .01 level.

The best predictor variable for acceptable rationale scores was comprehension scores; intelligence was also a significant predictor of acceptable rationale. Acceptable rationale scores were an extremely significant predictor of total cloze scores, however, comprehension was also found to be a significant predictor.

Next the test of multiple regression was conducted with the findings for the total grade three population to determine the best predictor for total cloze scores; the criterion variable being total cloze scores and the predictor variables being vocabulary scores, comprehension scores and intelligence quotients. The means and standard deviation of the variables are presented in table 4.24.

TABLE 4.24

MEAN, STANDARD DEVIATIONS OF VARIABLES
FOR TOTAL POPULATION--GRADE THREE

	Variable	M	S.D.
1.	Total Cloze Scores (T.C.)	17.26	6.2
2.	Vocabulary Scores (V)	4.27	.9
3.	Comprehension Scores (C)	4.14	1.2
4.	Intelligence Quotient (I.Q.)	104.35	11.4

For the total grade three population the mean score on the fifty-item cloze test was 17.26; the mean grade equivalent score on the vocabulary test was 4.27 and on the comprehension test was 4.1. The mean intelligence quotient was 104.35.

The correlation coefficients for the four variables are presented in table 4.25.

TABLE 4.25

CORRELATION COEFFICIENTS FOR VARIABLES--
GRADE THREE

	T.C.	V	C	I.Q.
T.C.	1.00			
V	.62	1.00		
C	.70	.69	1.00	
I.Q.	.49	.47	.58	1.00

A correlation of .70 between total cloze scores and comprehension was the highest correlation coefficient obtained for the total grade three population which was followed closely by a correlation of .69 between comprehension and vocabulary scores.

The correlation coefficients were converted into beta weights for the purpose of comparison in comparable units. The results are recorded in table 4.26. Also, the results of the T test are included in the tables.

TABLE 4.26

BETA WEIGHTS FOR PREDICTOR VARIABLES--
GRADE THREE

	Beta Weights for Predictors		
Criterion Variable	V	C	I.Q.
Total Cloze Score	.237	.478	.107
T	2.327*	4.347**	1.187 (n.s.)

*Significant at the .05 level.
**Significant at the .01 level.

In terms of the total grade three population comprehension scores were the best predictors of total cloze scores,

however, when only the vocabulary and intelligence variables were compared, vocabulary was a significant predictor of total cloze scores.

Next, the findings of the multiple regression test are reported as they relate to the grade five populations. First the interview population and then the total grade five population are examined.

The means and standard deviations of the variables for the interview population in grade five are reported in table 4.27.

TABLE 4.27

MEANS, STANDARD DEVIATIONS OF VARIABLES
FOR INTERVIEW POPULATION--GRADE FIVE

	Variable	M	S.D.
1.	Acceptable Rationale Scores (A)	14.00	5.0
2.	R1 Rationale Scores (R1)	5.00	3.0
3.	R2 Rationale Scores (R2)	1.367	2.4
4.	Total Cloze Scores (T.C.)	22.567	5.7
5.	Vocabulary Scores (V)	6.457	1.6
6.	Comprehension Scores (C)	5.997	2.1
7.	Intelligence Quotients (I.Q.)	108.167	16.0

The mean in total cloze scores for the grade five population was 22.567 with mean in acceptable rationale at 14. The grade equivalent mean score on the vocabulary test was 6.457 and the mean in comprehension was 5.997. Mean intelligence obtained was 108.167.

Correlation coefficients for the variables are presented in table 4.28.

TABLE 4.28

CORRELATION COEFFICIENTS FOR VARIABLES
FOR INTERVIEW POPULATION--GRADE FIVE

	A	R1	R2	T.C.	V	C	I.Q.
A	1.00						
R1	.06	1.00					
R2	-.13	-.55	1.00				
T.C.	.91	.14	.03	1.00			
V	.60	.21	.11	.64	1.00		
C	.64	.07	.22	.69	.64	1.00	
I.Q.	.78	-.08	.09	.81	.69	.74	1.00

A correlation of .91 between acceptable rationale and total cloze scores was the highest correlation coefficient obtained for the interview population in grade five. This was followed by a correlation of .81 between total cloze scores and intelligence. The next largest correlation was .78 between acceptable rationale and intelligence.

Conversion of correlation coefficients into beta weights is reported in table 4.29. Results of the test of significance are also included in this table.

Intelligence quotients were the best predictor variable for acceptable rationale scores and acceptable rationale was the best predictor for total cloze scores. Another significant predictor for total cloze scores was intelligence quotients.

Then the test of multiple regression was conducted with the findings for the total grade five population to determine the best predictor for total cloze scores. The criterion variable was total cloze scores with the predictor variables being

vocabulary scores, comprehension scores and intelligence quotients. The mean and standard deviations of the variables under consideration are presented in table 4.30.

TABLE 4.29

BETA WEIGHTS FOR PREDICTOR VARIABLES--
GRADE FIVE

	Beta Weights for Predictors		
Criterion Variable	A	R1	R2
Total Cloze Score	.936	.247	.292
T	14.732**	3.260**	3.831**
Criterion Variable	V	C	I.Q.
Total Cloze Score	.114	.175	.599
T	.712 (n.s.)	1.025 (n.s.)	3.310**
Criterion Variable	V	C	I.Q.
Acceptable Rationale Score	.094	.120	.628
T	.543 (n.s.)	.652 (n.s.)	3.221**

*Significant at the .05 level.
**Significant at the .01 level.

TABLE 4.30

MEAN, STANDARD DEVIATIONS OF VARIABLES
FOR TOTAL POPULATION--GRADE FIVE

	Variable	M	S.D.
1.	Total Cloze Scores (T.C.)	22.2	5.4
2.	Vocabulary Scores (V)	6.6	1.6
3.	Comprehension Scores (C)	5.9	2.0
4.	Intelligence Quotients (I.Q.)	106.3	13.1

The mean score for the fifty-item cloze test was 22.2 with the mean vocabulary for the total population being 6.6 and

5.9 for comprehension. The mean I.Q. was 106.3.

In table 4.31, the correlation coefficients for these variables are given.

TABLE 4.31

CORRELATION COEFFICIENTS OF VARIABLES
FOR TOTAL POPULATION--GRADE FIVE

	T.C.	V	C	I.Q.
T.C.	1.00			
V	.62	1.00		
C	.68	.72	1.00	
I.Q.	.71	.61	.76	1.00

The highest correlation for the total grade five population was .76 which was calculated between comprehension and intelligence; that was followed by a correlation of .72 between comprehension and vocabulary, which was closely followed by a correlation of .71 between total cloze scores and intelligence.

Table 4.32 presents the beta weights for the correlation coefficients for each set of compared predictor variables along with the results of the T test.

TABLE 4.32

BETA WEIGHTS FOR PREDICTOR VARIABLES--
GRADE FIVE

	Beta Weights for Predictors		
Criterion Variable	V	C	I.Q.
Total Cloze Score	.227	.204	.414
T	2.340*	1.719 (n.s.)	3.957**

*Significant at the .05 level.
**Significant at the .01 level.

The best predictor of total cloze scores for the grade five total population was intelligence quotients. Vocabulary was estimated to be the next best predictor.

Descriptive Analysis of Items

The items on the two cloze tests will be analysed in several ways. First, the items will be categorized according to grammatical class and then each category will be visually inspected to determine which grammatical classes of cloze items were most frequently replaced by exact words and which cloze items obtained acceptable rationale credit most frequently, and which obtained the other categories of rationale response credit.

Item on Grade Three Cloze Test

The cloze items were categorized into nouns, pronouns, verbs, adjectives, adverbs, conjunctions, prepositions, and definite articles. Table 4.33 presents the group of cloze categorized as nouns, pronouns, possessive pronouns, as well as a breakdown of the type of rationale obtained for these items.

This class of words contains a large number of the cloze items which were frequently filled by exact word replacements. For example, in item 36 _his_ was replaced 29 out of a possible 30 times with the exact word, with 24 out of the 29 subjects giving an acceptable rationale response. Other high frequency items were numbers 15 _sun_; 24 _day_; 30 _family_; 42 _tree_; and 51 _he_.

Table 4.34 presents the frequency totals for the articles, prepositions and conjunctions.

TABLE 4.33

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENTS FOR NOUNS, PRONOUNS,
POSSESSIVE PRONOUNS--GRADE THREE CLOZE TEST

Items	Frequency of Rationale Responses					Frequency of exact Word Replacement
	A	R1	R2	N	O	
4. things						0/30
12. day	12*	4	2	2	3	23/30**
15. sun	23*				2	25/30*
20. bed	15*	3			1	19/30
24. day	17*	3	1	1	1	23/30*
30. family	13*				1	14/30
31. snare						0/30
32. way	4	3		3	3	13/30
37. strand						0/30
42. tree	16*	1		1		18/30
47. time	2	6	1	1	1	11/30
48. sky	3	1			1	5/30
50. Che-Ka-Pash	1					1/30
27. he	16*	5			1	22/30**
14. us	1					1/30
36. his	24*	3		1	1	29/30**

*In top ten of acceptable rationale frequency.
**In top ten of exact word replacement frequency.

TABLE 4.34

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENTS FOR ARTICLES, PREPOSI-
TIONS, CONJUNCTIONS--GRADE THREE CLOZE TEST

Items	Frequency of Rationale Responses					Frequency of exact Word Replacement
	A	R1	R2	N	O	
7. the		2		3	2	7/30
13. the	1	6	3	4		14/30
18. the	1	11	5	6	4	27/30**
21. the		4	6	2	2	14/30
41. the		12	3	3	4	22/30**
33. about	4	5	3	2	2	16/30
29. to	4	5	1	7	3	20/30**
46. and	7	2	3	2	4	18/30
11. they		1				1/30
17. if	2	4	3	2	4	15/30

**In top ten of exact word replacement frequency.

In contrast to the preceding group, the frequency of acceptable rationale is very low for this category of words, however, it has in common with the previous group a high frequency of exact word replacements. Of particular interest are items number 18 the; 41 the; and 29 to, all of which were in the top ten in terms of frequency for exact word replacements.

Frequencies for verbs, adverbs and adjectives are presented in table 4.35.

Generally the frequencies are quite low for cloze items in the category of verb, adverb and adjectives with the exception of acceptable rationale frequencies for item number 2 always and number 19 not, which fell into the top ten and exact word replacement frequencies for number 34 thought and number 2 always which were also in the top ten items in terms of highest frequency of exact word replacements.

In summary, the group of words comprising of nouns, pronouns and possessive pronouns contained eight out of the top ten items for which acceptable rationale was most frequently given as well as having five of the top ten items for which exact word replacements were most frequently provided. Three of the cloze items in the articles, prepositional class were in the top ten in terms of frequency of exact word replacement but acceptable rationale frequencies were low with repetitions (R1) and "only one that fits" (R2) rationale responses predominating. Both frequencies of acceptable rationale and exact word replacements were low in the verb, adverb and adjective grouping with some exceptions noted.

TABLE 4.35

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENTS FOR VERBS, ADVERBS AND
ADJECTIVES--GRADE THREE CLOZE TEST

Items	Frequency of Rationale Responses					Frequency of exact Word Replacement
	A	R1	R2	N	O	
3. trying	1					1/30
5. fell	7	5				12/30
6. trying						0/30
8. didn't						0/30
22. watched	8	1	1		1	11/30
23. set	7				1	8/30
28. set	1					1/30
34. thought	8	6	1	1	4	20/30**
45. came		1	1			2/30
43. set	2			1		3/30
1. curious	1					1/30
10. father's						0/30
38. longest	6				2	8/30
39. next	6	4	2	1		13/30
44. topmost						0/30
9. another						0/30
2. always	9*	5	2	3	1	20/30**
25. as						0/30
35. so				2		2/30
40. before	2				1	3/30
26. then	5	4	1	5	2	17/30
19. not	9*	5	1	2	1	18/30
16. how	3	7	5	1	1	17/30

*In top ten of acceptable rationale frequency.
**In top ten of exact word replacement frequency.

Items on the Grade Five Cloze Test

Cloze items on the grade five test were categorized and frequency counts were made for acceptable rationale responses and exact word replacements. The findings for the nouns, pronouns, possessive pronouns are presented in table 4.36.

TABLE 4.36

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENTS FOR NOUNS AND PRO-
NOUNS--GRADE FIVE CLOZE TEST

Items	Frequency of Rationale Responses					Frequency of exact Word Replacement
	A	R1	R2	N	O	
5. century	1	4			1	6/30
6. street	1					1/30
9. care						0/30
15. cone	21*	2		1		24/30**
20. Marchiony	27*					27/30**
43. idea	1					1/30
45. rest	5					5/30
13. ice	26*			2		28/30**
26. ice	28*					28/30**
30. ice	27*	1		1		29/30**
33. ice	28*			1		29/30**
11. who		4	1	2	1	8/30
21. his						0/30
24. he	22*				1	23/30
31. his	1					1/30
32. his	1					1/30

*In top ten of acceptable rationale frequency.
**In top ten of exact word replacement frequency.

The nouns, pronoun grouping contains seven of the cloze items found in the top ten in terms of frequency of exact word

replacement and seven of the highest frequency items in terms
of acceptable rationale.

Conjunction, prepositions and article frequencies are
presented in table 4.37.

TABLE 4.37

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENTS FOR CONJUNCTIONS,
PREPOSITIONS AND ARTICLES--GRADE FIVE
CLOZE TEST

Items	Frequency of Rationale Responses					Frequency of exact Word Replacement
	A	R1	R2	N	O	
3. but	2	7	3	1	2	15/30
12. and		1	1	2	1	5/30
4. until						0/30
23. that		10	2	2	2	16/30
39. because	13*	7		1	2	23/30
2. of	1	7	8	5	3	24/30**
19. of	12	8	2	2		24/30**
35. of		10	1	1		12/30
41. in	20*	1	1	1		23/30
18. a	8	6		2	1	17/30
29. a	10	9	4	2		25/30**
48. the	3	9	2	2	3	19/30

*In top ten of acceptable rationale frequency.
**In top ten of exact word replacement frequency.

The same pattern as was evident with grade three items
is repeated with grade five cloze items. The conjunction, pre-
positions and article grouping containing a number of high fre-
quency items in terms of exact word replacement, however,
acceptable rationale totals are lower with a marked increase
in R1 and R2 type responses.

TABLE 4.38

FREQUENCY OF RATIONALE CATEGORIES AND EXACT
WORD REPLACEMENT FOR ADJECTIVES, ADVERBS
AND VERBS--GRADE FIVE CLOZE TEST

Items		Frequency of Rationale Responses					Frequency of exact Word Replacement
		A	R1	R2	N	O	
1.	has	2	4	3	6	2	17/30
2.	breaking	11			1	1	13/30
16.	was	3	19*	3	2	1	28/30**
34.	was	4	19*	2	3		28/30**
38.	were	5	12	2	1	1	21/30
46.	wasn't	2	2				4/30
49.	introduced	3					3/30
7.	glass	9	1				10/30
8.	great	19*					19/30
20.	new	13					13/30
25.	new	13					13/30
27.	thin						0/30
36.	edible						0/30
32.	popular	11				1	12/30
42.	toot	12					12/30
47.	world's						0/30
11.	away	9			3		12/30
14.	when	9	5	1			15/30
28.	up	10					10/30
44.	never						0/30
50.	so		1	2	1	1	4/30

*In top ten of acceptable rationale frequency.
**In top ten of exact word replacement frequency.

The verb, adverb, adjective grouping contains a very
small number of high frequency items as was the case with the
grade three cloze test items. Rationale responses of the R1 and
R2 type were much less predominant than was the case in the pre-
ceding grouping.

CHAPTER 5

SUMMARY, CONCLUSIONS AND IMPLICATIONS

The purpose of this study was twofold. The first objective was to obtain empirical data related to the ability to complete cloze items with exact word replacements and the ability to give acceptable rationale for cloze responses; and to determine the relationship of total cloze scores to reading ability, intelligence and sex. A second purpose of the investigation was to determine the best predictor variables for total cloze scores and for acceptable rationale scores.

The two main questions for study were:

1. What relationship exists between the ability to complete cloze items with exact word replacements and the ability to give an acceptable rationale for cloze responses at. grades three and five?

2. What relationship exists between cloze scores based on exact word replacement and measures of vocabulary, comprehension and intelligence as well as sex at grades three and five?

The ten hypotheses, which focused on these two main areas of investigation, were grouped for purposes of summarizing and and reviewing findings.

Hypotheses One and Two

The first area of investigation involved the relationship of total cloze scores to acceptable rationale scores at

grades three and five. These were examined under hypotheses
one and two. A comparison was made between low and high cloze
scorers at each grade based on the subject's performance on a
cloze test and his/her responses during individual interviews
in which rationale for cloze items was requested.

Hypotheses Three Through Ten

The remaining hypotheses focused on the second main
question of the study, which was the relationship of scores
obtained on cloze tests based on basal reader material at each
grade's grade placement level to grade equivalent scores ob-
tained on Gates-MacGinitie vocabulary and comprehension sub-
tests as well as to intelligence quotients obtained on the
Otis-Lennon Mental Ability Tests. In addition the relationship
of cloze scores to sex was examined.

This chapter summarizes the findings related to the
hypotheses, followed by a summary of the descriptive analysis
of the strategies and clues used, as well as a summary of the
item analysis. Conclusions and implications for classroom
practice and further research end the chapter.

Summary

Total precloze scores were obtained from 91 grade threes
and 99 grade fives. Cloze scores obtained for each grade were
ranked from lowest to highest and a median was calculated. Then
using a random numbers table, 15 of the lowest and 15 of the
highest cloze scorers for each grade were selected for indivi-
dual interviews. In these 60 interviews, rationale was requested

for each of the 50 cloze items for which a response was made,
however, only the rationale responses for exact word replace-
ments were evaluated and classified in one of five possible
categories: A--acceptable; R1--repetition; R2--"only one that
fits"; N--not acceptable; O--omitted, no rationale given. Each
subject's total acceptable rationale score was then paired with
his/her total cloze score. A comparison was made between low
and high cloze scorers at grades three and five.

Next a comparison of the total cloze scores, of low and
high vocabulary scorers and low and high comprehension scorers
based on grade equivalent scores on the Gates-MacGinitie sub-
tests and of low and high intelligence quotient scorers based
on intelligence quotients obtained from the Otis Lennon Mental
Ability Test, were computed. Scores obtained on each of these
three measures were ranked and medians were calculated before
comparison could be made. Finally total cloze scores of boys
and girls were compared.

Summary of Findings Related to
Hypotheses

1. There were significant differences in the acceptable
rationale scores of low and high cloze scorers in grade three
which favored the high cloze scorers.

2. A significant difference was found in the acceptable
rationale scores of low and high cloze scorers in grade five when
the Wilcoxon, Mann, Whitney Test was applied, however, when the
more powerful statistical analysis of co-variance controlled for
initial difference in total cloze scores, no significant differ-
ence emerged.

3. Low and high vocabulary scorers in both grades three and five were found to differ significantly in terms of total cloze scores, with high vocabulary scorers having the advantage over low vocabulary scorers.

4. A significant difference in total cloze scores was found betwen low and high comprehension scorers in both grades three and five, with high comprehension scorers having the advantage over low comprehension scorers.

5. Subjects at both grades three and five with high intelligence quotients were found to have a significant advantage over subjects with low intelligence quotients in terms of total cloze scores.

6. No significant differences were found between boys and girls at either grade three or five in terms of total cloze scores.

Summary of Descriptive Analysis

1. The mean total cloze scores of the grade three interview population was 17.6 and the mean acceptable rationale score was 8.2. In contrast, the mean total cloze score for the grade five interview population was 27.5 and the mean acceptable rationale score was 14.0.

2. Mean grade equivalent scores in vocabulary and comprehension for the total sample in each grade were very near the grade placement level of the subjects. Vocabulary mean was 4.27 and comprehension mean was 4.14 for grade three with grade placement level of 3.9; vocabulary mean was 6.6 and comprehension mean was 5.9 for grade five with grade placement level of 5.9.

3. Mean intelligence quotient for grade three was 104.35 and 106.3 for grade five.

4. With the grade three interview population, it was found that comprehension was the best predictor of acceptable rationale scores and acceptable rationale scores were the best predictor of total cloze scores.

5. With the total grade three population, comprehension was identified as the best predictor of total cloze scores.

6. With the grade five interview population, intelligence quotients were found to be the best predictors of acceptable rationale scores and acceptable rationale scores were the best predictors of total cloze scores.

7. With the total grade five population, intelligence quotients were the best predictors of total cloze scores.

8. Subjects at both the grades three and five made use of a variety of strategies when engaging in the selection of words for cloze items ranging from use of word classifications, immediate and distant context, retroactive context, as well as trying out absurd or illogical words to verify their choices. Finally, subjects employed the technique of repeating sentences and phrases containing the cloze item, all the while emphasizing the cloze response very emphatically to rationalize the appropriateness of their choice.

9. Subjects used a variety of reasons when providing rationale for cloze items. The reasons ranging from grammatical reasons including reference to pronouns, possessive pronouns, past tense, predicate verbs, noun determiners, conjunctions,

abbreviations, prepositions, punctuations; to semantical clues including reference to main idea, story detail, story characters, as well as use of own experiential background, own logic and knowledge of expressions. Syntactical clues were also used including the subject's feel for word order or juxtaposition of words, sentence structure and word function.

10. Items most frequently replaced by exact words were nouns, proper nouns, pronouns, possessive and relative pronouns, noun determiners and common prepositions. A small number of verbs and adverbs were included in the high frequency range.

11. Classification of words for which acceptable rationale was most frequently obtained were nouns, proper nouns, possessive pronouns, pronouns. Two adverbs, one adjective, one conjunction and one preposition were also included in the high frequency range.

12. Classes of words for which rationale responses obtained fell into the R1--repetition, or R2--"only one that fits" categories, were mostly the conjunctions, prepositions, article and predicate verbs, class of words.

Conclusions

The following conclusions are offered, bearing in mind the danger of generalizing from a single study.

1. Low cloze scorers in grade three are significantly poorer at providing rationale for their cloze responses than are high cloze scorers. Low cloze scorers in grade five were not significantly poorer than their counterparts in providing rationale for their cloze responses. This contrasting finding

would seem to indicate that there is a variation in the rea-
soning powers of grade threes whereas by grade five, the rea-
soning strength of the good readers has levelled out and is at
par with the poor readers. This is in contradiction to Hafner's[1]
finding with college students. He concluded that poor readers
at the college level are less able to reason, and less able to
use contextual clues than were good readers.

One explanation for this difference between the grade
threes and fives could be that overall, the grade five cloze
test was an easier test for the grade fives than was the grade
three cloze test for the grade threes in spite of the fact that
the readability of each cloze test was at the respective grade
placement at the time of administration. This was evidenced
not only in total cloze score means which were 17.6 for grade
three and 27.5 for grade five, but also by visual inspection of
sentence structure and sentence length. The average grade three
sentence was 23.2 words long, whereas the average sentence length
on the grade five test was 19.6 words in length.[2] Another factor
to be considered was that the grade five test had ten words that
were considered "difficult" whereas the grade three test had
only four "difficult" words. It should be noted, however, that
one word in the grade three passage not included in these iden-
tified difficult words was the word snare, a concept which most

---

[1]Lawrence E. Hafner, "Relationship of Various Measures
to the Cloze," New Concepts in College-Adult Reading, ed. Eric
L. Thurston and L. E. Hafner, Thirteenth Yearbook of the
National Reading Conference, 1964, p. 138.

[2]See Appendix B.

grade threes in the interview population did not know. This word was crucial to the understanding of the grade three passage, whereas none of the words found difficult by grade fives were as critical to the understanding of their passage. Further, the four cloze items required the response _ice_ and these four items were particularly easy in terms of giving acceptable rationale.

Another factor to be considered is that the main idea of the grade three passage required inference ability whereas the grade five passage was more factual and straightforward.

In conclusion, it could be said that if the cloze test for grade five had been comparable to the grade three test, the findings would have been almost the same for each grade and would have been in keeping with Hafner's findings with college students.

2. High vocabulary and high comprehension scorers complete cloze items more frequently with exact word replacements than do low vocabulary and low comprehension scorers. This finding supports the results of Hafner,[1] Jenkinson,[2] Rankin,[3]

---

[1]Lawrence E. Hafner, "Relationship of Various Measures to the Cloze," in Thirteenth Yearbook of the National Reading Conference, ed. E. L. Thurston and L. E. Hafner, Milwaukee, Wisconsin: The National Reading Conference, 1963, p. 144.

[2]Marion D. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension," (Doctoral dissertation abstract, University of Chicago, 1971).

[3]Earl F. Rankin, "An Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension," (Doctoral dissertation abstract, University of Michigan, 1958).

Ruddell,[1] and Friedman.[2] These studies, which were mostly done with more mature pupils, obtained high to moderately high correlations between cloze test scores and measures of comprehension and/or vocabulary. This study confirms Gallant[3] and Schneyer's[4] findings with primary and elementary pupils.

3. Subjects with high intelligence quotients are more successful at completing cloze items with exact word replacements than subjects with low intelligence quotients. Intelligence and the ability to reason appear to play an important role in a subject's ability to cope with cloze tasks. This confirms the findings of Hafner,[5] Taylor,[6] Jenkinson,[7]

[1]Robert B. Ruddell, "A Study of the Cloze Comprehension Technique in Relation to Structurally Controlled Reading Materials," in Improvement of Reading Through Classroom Practice, ed. J. A. Figural (Newark, Delaware: International Reading Association, 1964).

[2]M. Friedman, "The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at the University of Florida (Doctoral dissertation abstract, University of Florida, 1964).

[3]Ruth Gallant, "Use of Cloze Tests as a Measure of Readability in the Primary Grades," in Reading and Inquiry, ed. J. A. Figurel (Newark, Delaware: International Reading Association, 1965): 286-87.

[4]W. J. Schneyer, "Use of the Cloze Procedure in Improving Reading Comprehension," The Reading Teacher 30 (December 1965): 174-79.

[5]Lawrence E. Hafner, "Relationship of Various Measures to the Cloze," in Thirteenth Yearbook of the National Reading Conference, ed. E. L. Thurston and L. E. Hafner, Milwaukee, Wisconsin: The National Reading Conference, 1963, p. 144.

[6]W. L. Taylor, "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," Journal of Applied Psychology 41 (February 1957): 19-26.

[7]Marion D. Jenkinson, "Selected Processes and Difficulties in Reading Comprehension," (Doctoral dissertation abstract, University of Chicago, 1971).

Fletcher[1] and Schneyer[2]. They found that intelligence correlated high to moderately high with cloze scores. Hafner found that reasoning as revealed by connective and incongruency scores demonstrated the importance of reasoning in the cloze procedure.

    4. No differences were found between boys and girls regarding their ability to complete cloze items. This finding is in opposition to most reading research which generally favors girls. This difference may be due to the fact that the population used was possibly not a representative sample in that all the pupils from one school were used with no attempt made to randomize the selections. The uneven boy/girl ratios possibly affected the results also, with 38 girls compared to 53 boys in grade three, and 54 girls versus 45 boys in grade five.

    5. For grade three, comprehension scores are the best predictor of total cloze scores whereas for grade five, intelligence quotients are the best predictors of cloze scores. This difference could be the result of the difference in the format of the intelligence tests used. The Otis-Lennon used with the grade threes required no reading being an entirely oral test dealing with pictures and designs whereas the Otis-Lennon for

---

[1] J. E. Fletcher, "A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities," University of Washington, 1959.

[2] W. J. Schneyer, "Use of the Cloze Procedure in Improving Reading Comprehension," The Reading Teacher 30 (December 1965): 174-79.

the grade five subjects was a written test requiring silent
reading.

6. Grade threes who scored highest on comprehension
tests were the pupils best able to give acceptable reasons for
their cloze responses, revealing an adequate understanding of
the logic and relationship of ideas while reading and the abil-
ity to express this understanding. Grade fives who scored
highest on intelligence were the pupils best able to give ac-
ceptable reasons for their cloze response. In view of the
silent reading format of the particular intelligence test, it
can be concluded that the type of measurement which tests the
pupil's ability to reason and recognize logical relationships
in a reading context will also best predict acceptable rationale
scores. Because the grade three intelligence test was more a
listening test and did not deal directly with written language,
it probably did not correlate as highly with total cloze scores.

7. Pupils are more successful in replacing nouns, pro-
per nouns, pronouns, possessive pronouns and noun determiners,
some common prepositions and some conjunctions in cloze tests
than words in other categories of speech such as verbs, adverbs,
and adjectives.

8. Pupils are more successful in supplying acceptable
rationale for nouns, proper nouns, pronouns, possessive pronouns
and some occasional verbs and adjectives. Pupils found it dif-
ficult to give their reasons for selecting the more abstract
words such as noun determiners, conjunctions, prepositions, pre-
dicate verbs, and adverbs.

9. Pupils at these early grades are able to cope with this rather difficult and abstract task of verbalizing their reasons for cloze items. By their responses in the individual interviews, pupils have demonstrated that they can engage in this introspective and retrospective process quite well and that they can use a wide variety of linguistic cues to substantiate their choices.

10. Pupils at grade three and five employ a wide range of strategies and clues when completing cloze tasks. A conclusion could be made that they engage in this kind of processing when reading at any time. If so, then those pupils who are more adept at using strategies and giving reasons are more successful in other reading activities and would confirm the theory that reading is a thinking process involving an interrelation of ideas and language.

11. There were no major differences in the ability of grade threes and fives in terms of using strategies and clues when processing cloze items excepting that grade fives began to use labels, such as "conjunction," "article," "consonant," and "past tense" when giving their rationale responses, whereas grade threes tended to remark, "it sounded best" or repeated the phrase continuing the item, in an attempt to explain their intuitive knowledge of language.

Implications for the Classroom

The following implications are apparent on the basis of this present investigation:

1. Cloze tests may be used with a fair degree of

confidence when administered in conjunction with other reading achievement tests. Since cloze test results tend to correlate moderately high with other reading measures, a cloze test becomes a check against those standardized tests which involve multiple choice items; such items becoming very susceptible to guessing or random marking of answers. In contrast to most standardized reading tests, cloze tests measure the reader's ability to cope with complex sentence structure and concepts in the reading material. Thus, a combination of cloze and standard reading tests is recommended for use in the classroom.

2. If rationale is requested resulting from the completion of a cloze test in the manner in which it was obtained in this study, then the cloze test may be used as a diagnostic tool, to determine what cues and strategies a particular reader is able to use in a reading situation. Additional information about a reader could be gleaned from an analysis of the item responses. Readers who supply synonyms or semantically acceptable responses show more strength and demonstrate a thinking approach to reading as contrasted to the reader who supplied totally unacceptable responses, ignoring thought units, punctuation and logic.

3. It would seem advisable, that since reasoning and awareness of the interrelationship of ideas play such an important role in reading, and since the cloze procedure so uniquely accommodates this kind of processing, that time by spent in the classroom doing cloze tasks first in oral form and in small group situations, so as to expose the young readers to the logic

and interaction of ideas and language, and to help them to see reading as a "psycholinguistic guessing game" rather than the mere perception of word images.

Implications for Further Research

The following suggestions for further research are proposed in view of the results of the present study:

1. The rationale-requesting portion of this study should be expanded to cover a wider age/grade range and with more detailed and thorough analysis of rationale responses in an effort to compile a continuum of strategies and reasons across the grades and/or ages.

2. The rationale for the incorrect cloze responses contains a wealth of information regarding the logic and reasoning processes of readers when faulty assumptions and conclusions are made. These rationale responses should be analyzed to gain further insight into the differences between good and poor readers reasoning ability.

3. It would appear that R1 and R2 rationale responses are basically the same and should be joined to form one response category. This conclusion is based on the fact that many subjects used these two responses interchangeably or in conjunction with one another.

4. Rather than a single cloze test, two cloze tests should be given to each subject: one common cloze test across the grades, or at least across a range of two or three grades; and one at the grade placement level of the subjects at the time of the testing to allow across grade comparisons on a standard

cloze test and comparisons on a cloze test gaged to the expected reading level of the subjects.

5.   A more detailed analysis of cloze items should be done to determine at what grades particular classes of words are replaced more frequently with exact words, and at what grades are acceptable rationales obtained for particular classes of words.   This would aid in tracing the development of a child's understanding of the many linguistic cues inherent in written material.

SELECTED BIBLIOGRAPHY

Books

Dawkins, John. Syntax and Readability. Newark, Delaware: Inter-
     national Reading Association, 1975.

Jongsma, E.   The Cloze Procedure as a Teaching Technique. Newark,
     Delaware: International Reading Association, 1971.

Robinson, R. D.   An Introduction to the Cloze Procedure--An
     Annotated Bibliography.   Newark, Delaware: International
     Reading Association, 1972.

Tests

Gates, Arthur I., and MacGinitie, Walter H.   Gates-MacGinitie
     Reading Tests.   Columbia University: Teachers College
     Press, 1965.

Otis, Arthur S., and Lennon, Roger T.   Otis-Lennon Mental Ability
     Test.   New York: Harcourt, Brace and World, Inc., 1967.

Articles and Periodicals

Ames, Wilbur S.   "The Development of a Classification Scheme of
     Contextual Aids."   The Reading Research Quarterly 1 (Fall
     1966): 57-82.

Bickley, A. C.; Ellington, Billie J.; and Bickley, R. T.   "The
     Cloze Procedure: A Conspectus."   Journal of Reading
     Behavior 2 (Summer 1970): 232-49.

Bloomer, R. H.   "The Cloze Procedure as a Remedial Reading Exer-
     cise."   Journal of Developmental Reading 5 (Spring 1962):
     173-81.

Bormuth, John R.   "Cloze as a Measure of Readability."   Reading
     as an Intellectual Activity, ed. J. A. Figurel, Inter-
     national Reading Association Conference Proceedings 8
     (1963): 131-34.

Bormuth, John R.   "Experimental Applications of Cloze Tests."
     Improvement of Reading Through Classroom Practice, ed.
     J. A. Figurel, International Reading Association Con-
     ference Proceedings 9 (1964): 303-6.

Bormuth, John R. "Validities of Grammatical and Semantic Clas-
    sifications of Cloze Test Scores." Reading and Inquiry,
    ed. J. A. Figurel, International Reading Association
    Conference Proceedings 10 (1965): 283-85.

Bormuth, John R. "Optimum Sample Size and Cloze Test Length
    in Readability Measurement." Journal of Educational
    Measurement 2 (June 1965): 111-16. .

Bormuth, John R. "Comparable Cloze and Multiple-Choice Compre-
    hension Test Scores." Journal of Reading 10 (February
    1967): 291-99.

Bormuth, John R. "Factor Validity of Cloze Tests as Measures
    of Reading Comprehension." The Reading Research
    Quarterly 3 (Spring 1969): 358-62.

Bormuth, John R., and MacDonald, O. L. "Cloze Tests as a Measure
    of Ability to Detect Literary Style." Reading and
    Inquiry, ed. J. A. Figurel, International Reading Asso-
    ciation Conference Proceedings 10 (1965): 287-90.

Buros, Oscar K. The Third Mental Measurements Yearbook. New
    Brunswick, N.J.: Rutgers University Press, 1949.

Buros, Oscar K. The Seventh Mental Measurements Yearbook.
    Highland Park, N.J.: The Gryphon Press, 1972.

Culhane, Joseph W. "Cloze Procedures and Comprehension." The
    Reading Teacher 23 (February 1970): 411-14.

Gallant, Ruth. "Use of Cloze Tests as a Measure of Readability
    in the Primary Grades." Reading and Inquiry, ed. J. A.
    Figurel, International Reading Association Conference
    Proceedings 10, (1965): 286-87.

Goodman, Kenneth S. "Reading: A Psycholinguistic Guessing
    Game." Journal of the Reading Specialist (May 1967):
    126-35.

Gove, Mary K. "Using the Cloze Procedure in a First Grade Class-
    room." The Reading Teacher 29 (October 1975): 36-38.

Greene, Frank P. "Modification of the Cloze Procedure and
    Changes in Reading Test Performance." Journal of
    Educational Measurement 2 (December 1965): 213-17.

Guice, Billy M. "The Use of the Cloze Procedure for Improving
    Reading Comprehension in College Students." Journal of
    Reading Behavior 1 (Summer 1969): 81-90.

Hafner, Lawrence E. "Relationships of Various Measures to the
    Cloze." New Concepts in College-Adult Reading, ed. E. L.
    Thurston and L. E. Hafner, Thirteenth Yearbook of the
    National Reading Conference (1964): 135-45.

Hafner, Lawrence E. "Implications of the Cloze." The Philos- ophical and Sociological Bases for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference (1965): 151-58.

Hafner, Lawrence E. "Cloze Procedure." Journal of Reading 9 (May 1966): 415-21.

Heitzman, A., and Bloomer, R. "The Effect of Non-Overt Rein- forced Cloze Procedure on Reading Comprehension." Journal of Reading (December 1967): 213-23.

Jones, Margaret B., and Pikulski, Edna C. "Cloze for the Classroom." Journal of Reading 17 (March 1974): 432-38.

Kennedy, Delores K., and Weener, Paul. "Visual and Auditory Training with the Cloze Procedure to Improve Reading and Listening Comprehension." Reading Research Quar- terly 8 (Summer 1973): 527-41.

Kirby, Clara L. "Using the Cloze Procedure as a Testing Tech- nique." Reading Diagnosis and Evaluation, ed. D. L. DeBoer. Newark, Delaware: International Reading Asso- ciation (1970): 68-77.

Louthan, Vincent. "Some Systematic Grammatical Deletions and Their Effects on Reading Comprehension." English Journal 54 (April 1965): 295-99.

MacGinitie, Walter H. "Contextual Constraint in English Prose Paragraphs." Journal of Psychology 51 (January 1961): 128-31.

Mork, Theodore A. "Clozing the Placement Gap: A New Tool for Administerators and Teachers." Educational Leadership 28 (April 1971): 763-67.

Rankin, Earl F. "Uses of the Cloze Procedure in the Reading Clinic." Reading in a Changing Society, ed. J. A. Figurel. New York: Scholastic Magazines (1959): 262-63.

Rankin, Earl F. "Cloze Procedure--A Survey of Research." Philosophical and Sociological Bases for Reading, ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference (1965): 133-50.

Rankin, Earl F. "Grade Level Interpretation of Cloze Readability Scores." Reading: The Right to Participate, ed. F. P. Greene. Milwaukee, Wisconsin: The National Reading Conference (1971): 30-37.

Rankin, Earl F., and Culhane J. "Comparable Cloze and Multiple- Choice Comprehension Test Scores." Journal of Reading 10 (February 1967): 291-99.

Rankin, Earl F., and Overholser, Betty. "Reactions of Inter-
        mediate Grade Children to Contextual Clues." Journal
        of Reading Behavior 20 (March 1969): 50-73.

Ransom, Peggy E. "Determining Reading Levels of Elementary
        School Children by Cloze Testing." Forging Ahead in
        Reading, ed. J. A. Figurel, International Reading Asso-
        ciation Conference Proceedings 12 (1968): 477-82.

Ruddell, Robert B. "A Study of the Cloze Comprehension Technique
        in Relation to Structurally Controlled Reading Material."
        Improvement of Reading Through Classroom Practice, ed.
        J. A. Figurel, International Reading Association Confer-
        ence Proceedings 8 (1964): 298-303.

Ryan, Ellen B., and Semmel, Melvyn I. "Reading as a Construc-
        tive Language Process." Reading Research Quarterly 5
        (Fall 1969): 61-83.

Schneyer, Wesley J. "Use of the Cloze Procedure for Improving
        Reading Comprehension." The Reading Teacher 30 (December
        1965): 174-79.

Taylor, W. L. "Cloze Procedure: A New Tool for Measuring Read-
        ability." Journalism Quarterly 30 (Fall 1953): 415-38.

Taylor, W. L. "Recent Developments in the Use of 'Cloze Proce-
        dure'." Journalism Quarterly 33 (1956): 42-48.

Taylor, W. L. "'Cloze' Readability Scores as Indices of Indiv-
        idual Differences in Comprehension and Aptitudes."
        Journal of Applied Psychology 41 (February 1957): 19-26.

Weaver, Wendall W. "Theoretical Aspects of the Cloze Procedure."
        The Philosophical and Sociological Bases for Reading,
        ed. E. L. Thurston and L. E. Hafner, The National Reading
        Conference (1965): 115-32.

Weaver, Wendall W., and Kingston, Albert J. "A Factor Analysis
        of the Cloze Procedure and Other Measures in Reading
        and Language Ability." The Journal of Communication 12
        (December 1963): 252-61

Weber, Rose-Marie. "A Linguistic Analysis of First-Grade Reading
        Errors." Reading Research Quarterly 5.

Educational Resources Information Center Documents

Fram, Ralph D. "A Review of the Literature Related to the Cloze
        Procedure." ERIC Document ED 075 785, 1972.

Hibbard, Mike. "The Relationship Between Concept Attainment and
        Verbal Predictive Ability for Primary Grade Children."
        ERIC Document ED 084 100, 1973.

Unpublished Materials

Fletcher, J. E. "A Study of the Relationship Between Ability to Use Context Aid in Reading and Other Verbal Abilities." Doctoral dissertation, University of Washington, 1959.

Friedman, M. "The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at the University of Florida." Doctoral dissertation, University of Florida, 1964.

Horton, R. J. "The Construct Validity of Cloze Procedure: An Exploratory Factor Analysis of Cloze, Paragraph Reading and Structure of Intellect." Doctoral dissertation, Hofstra University, 1973.

Jenkinson, M. E. "Selected Processes and Difficulties in Reading Comprehension." Doctoral dissertation, University of Chicago, 1957.

Kirchoff, Leo H. "A Study Utilizing the Cloze Test to Determine Reading Levels of First Grade Children Who Have Been Taught Beginning Reading by Four Different Approaches." Doctoral dissertation, University of Kansas, 1968.

Rankin, Earl F. "An Evaluation of the Cloze Procedure As a Technique for Measuring Reading Comprehension." Doctoral dissertation abstract, University of Michigan, 1958.

Ruddell, Robert B. "An Investigation of the Effect of the Similarity of Oral and Written Patterns of Language Structure on Reading Comprehension." Doctoral dissertation, Indiana University, 1963.

APPENDIX A

CLOZE TESTS

PLEASE PRINT OR WRITE CLEARLY:

Name _____   Boy _____   Girl _____

Birthdate _____

Grade _____   Teacher _____   School _____

DIRECTIONS:

This is a new type of reading story to see how well you understand what
you read. You are to guess what words have been left out. Read this
sentence and think of a word that would complete it.

    We are going to _____ circus.

Does your word make the sentence complete?

Sometimes you will be able to think of a number of words that would fit
in the sentence but you are to pick only the one word that fits the
sentence best.

Now try to fill in the blanks in this paragraph.

    Mother Duck once made her nest beside the wall of a great palace.
A stream ran along ___1___ outside of the wall ___2___ here
the ducks and ___3___ ate and swam about. ___4___ Duck found it
very ___5___ sitting on her eggs ___6___ after day. She did
___7___ have many visitors. Her ___8___ liked better to swim
___9___ in the stream than ___10___ sit and talk to ___11___ .

As you can see all the blanks are the same size but the words you write
in the blanks may be short or long. Only one word can be used in each
blank. Do your best to fill in every blank. You will find some blanks
very easy and others hard. Do not be afraid to guess what a word might
be. You may find it best to leave some hard blanks until you have
completed the rest of the story and then come back and fill in those at
the end. If you don't know how to spell a word just try to spell it
the best you can and please print or write clearly. Now turn the page
and begin.

GRADE THREE

How Che-Ka-Pash Made the Sun Stand Still

Once upon a time, many, many years ago, in an Indian village far up in the north, there lived a little Indian boy called Che-Ka-Pash.   Now, Che-Ka-Pash was a _CURIOUS_ little boy and was _ALWAYS_ getting himself into trouble _TRYING_ to find out why _THINGS_ were so.
   1   2
        3   4

Once he _FELL_ into a cold river, _TRYING_ to find out why _THE_ falls
          5   6   7
near his village _DIDN'T_ freeze over in winter.   _ANOTHER_ time he broke his
                   8   9
FATHER'S bow trying to see _WHY_ it was springy.
   10   11

One _DAY_ , in what would be _THE_ month of March to _US_ , he sat
      12   13   14
watching the _SUN_ rise, and he thought, "_HOW_ nice it would be _IF_ it
               15   16   17
were daylight all _THE_ time. Then I would _NOT_ have to go to _BED_ early
                    18   19   20
at all. In _THE_ days that followed, Che-Ka-Pash _WATCHED_ the sun rise and
            21   22
SET , set and rise. Each _DAY_ got a little longer _AS_ the spring   .
 23   24   25
came. And _THEN_ Che-Ka-Pash had an idea. _HE_ had seen his father
            26   27
SET snares in the bush _TO_ catch food for his _FAMILY_ . Why not try
 28   29   30
to _SNARE_ the sun the same _WAY_ ? The more he thought _ABOUT_ it, the
     31   32   33 .
more he _THOUGHT_ it might be done.
          34

_SO_ one night he asked _HIS_ sister for a long _STRAND_ of her
 35   36   37
hair, the _LONGEST_ she could find. The _NEXT_ morning he was up _BEFORE_
            38   39   40
the sun. Climbing to _THE_ top of the tallest _TREE_ he could find, he
                      41   42
SET a snare on the _TOPMOST_ branch. Then he quickly _CAME_ down the
 43   44   45
tree again _AND_ waited.
            46

In a short _TIME_ the sun came up.   _AS_ it climbed into the _SKY_
            47   48   49
and neared his snare, _CHE-KA-PASH_ held his breath. "If I catch the sun in
                         50
my snare," he said to himself, "It will stay daylight until I let it go again."

GRADE FIVE

History of the Ice Cream Cone

Bit by bite, lick by lick, Americans consume more than four billion ice cream cones each year.

The ice cream cone __HAS__ become a taken-for granted part __OF__ every-
day life in America. __BUT__ it wasn't always so. __UNTIL__ the turn of the
__CENTURY__ ice cream sold by __STREET__ vendors was served in __GLASS__ dishes,
which required a __GREAT__ deal of washing and __CARE__ . And customers were
always __BREAKING__ the glasses or walking __AWAY__ with them. This
unsatisfactory __AND__ expensive means of serving __ICE__ cream changed in 1896
__WHEN__ the first ice cream __CONE__ was invented. The inventor __WAS__ an
Italian immigrant, Marchiony, __WHO__ sold ice cream from __A__ pushcart in
the streets __OF__ New York City.

Mr. __MARCHIONY__ got tired of washing __HIS__ serving dishes and buying
__NEW__ ones to replace those __THAT__ had been broken so __HE__ decided to
make a __NEW__ kind of container for __ICE__ cream. He baked a __THIN__ cooky and
rolled it __UP__ into the shape of __A__ cone. The he scooped __ICE__ cream into
the cone. __HIS__ customers were delighted with __HIS__ new way of serving
__ICE__ cream, and Mr. Marchiony __WAS__ glad to be free __OF__ dishes and dish-
washing. These __EDIBLE__ ice cream containers became __POPULAR__ in New York.
They __WERE__ called "toot" cones, probably __BECAUSE__ they resembled little
horns. __MR.__ Marchiony made a fortune __IN__ New York with his __TOOT__ cones,
but somehow the __IDEA__ of the toot cone __NEVER__ "caught on" in the __REST__
of the country. It __WASN'T__ until the St. Louis __WORLD'S__ Fair in 1904 that
__THE__ ice cream cone was __INTRODUCED__ to the world. It __SO__ happened that
an ice cream vendor at the Fair was enjoying a brisk business selling dishes
of ice cream for 5 cents and 10 cents. One day he ran out of dishes for ice
cream and didn't know what to do.

APPENDIX B

TIMETABLES

TESTING TIMETABLES

	GRADE THREE		GRADE FIVE	
Cloze Test	June 2	1:00 Class K.	June 13	9:00 Class P.
	June 3	9:15 Class G.		10:00 Class K.
		10:15 Class E.		11:00 Class S.
		11:00 Class P.		1:00 Class J.
Gates MacGinitie	(Primary C)		(Survey D)	
	June 9	1:00 Class K.	June 17	10:00 Class P.
	June 10	9:00 Class G.		11:00 Class K.
		10:00 Class E.		1:00 Class S.
		11:00 Class P.		2:00 Class J.
Otis-Lennon	Elementary I		Elementary II	
	(Form J)		(Form J)	
	June 16	9:00 Class G.	June 20	9:00 Class P.
		10:00 Class E.		10:00 Class K.
		11:00 Class P.		11:00 Class S.
		1:00 Class K.		1:00 Class J.

READABILITY

How Che-Ka-Pash Made the Sun Stand Still

Total number of words	116
Number of Sentences	5
Number of words not on (curious, springy, rise, freeze)	4
Average sentence length	23.2
4 x .121	2.80
3 x .082	.328
Constant	.659
Estimated Grade level	3.787

History of the Ice Cream Cone

Total number of words	118
Number of sentences	6
Number of words not on (consume, billion,granted, century, vendors, required, expensive, invented, immigrant, inventor)	10
Average sentence length	19.66
Dale Score	8.4
19.66 x .0496	.975
8.4 x .1579	1.326
Constant	3.6365
Raw score	5.937