DOCUMENT RESUME

ED 129 878   TM 005 680

AUTHOR   O'Reilly Robert P.; Streeter, Ronald E.
TITLE   Reports on the Development and Validation of a System
                for Measuring Literal Comprehension in a
                Multiple-Choice Cloze Format: Preliminary Factor
                Analytic Results.
PUB DATE   Mar 76
NOTE   33p.; Paper presented at the Annual Meeting of the
                American Educational Research Association (60th, San
                Francisco, California, April 19-23, 1976)

EDRS PRICE   MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS   *Cloze Procedure; Elementary Education; *Factor
                Analysis; *Multiple Choice Tests; *Reading
                Comprehension; Reading Tests; *Test Validity

ABSTRACT
                The results of a series of factor analyses of a new
test of literal comprehension using a multiple-choice cloze format
are summarized. These analyses were conducted in the validation of a
test design to measure for the most part a factor of literal
comprehension independent of IQ and inferential reading processes,
yet marked by certain related types of test items included in
standardized and other measures of literal comprehension. In this
study, the Multiple-Choice Cloze (MCC) test was administered to a
sample of 3,125 students in grades one to six in a medium-sized urban
school district in conjunction with its annual standardized testing
program. Besides the MCC, other measures included in the analyses
were an alternate measure of literal comprehension based on Bormuth's
wh-item, a measure of passage independence based on wh-items, the
Short Form Test of Academic Aptitude, and the California Achievement
Test. The factor analytic results support the conclusion that the MCC
measures literal comprehension, a trait that is essentially
independent of IQ. However, it was also determined that the MCC had
minor loadings on a second, and possibly a third, component related
to IQ, inferential reading skills, and language mechanics.
(Author)

Session 13.04

Reports on the Development and Validation of a System for Measuring
Literal Comprehension in a Multiple-Choice Cloze Format:

Preliminary Factor Analytic Results[*]

Robert P. O'Reilly and Ronald E. Streeter
March 1976

Bureau of School and Cultural Research
New York State Education Department
Albany, New York

ACKNOWLEDGMENT

Abstract

This report summarizes the results of a series of factor analyses of a new test of literal comprehension using a multiple-choice cloze format. These analyses were conducted in the validation of a test designed to measure for the most part a factor of literal comprehension independent of IQ and inferential reading processes, yet marked by certain related types of test items included in standardized and other measures of literal comprehension. In this study, the Multiple-Choice Cloze (MCC) test was administered to a sample of 3,125 students in grades 1-6 in a medium-sized urban school district in conjunction with its annual standardized testing program. Besides the MCC, other measures included in the analyses were an alternate measure of literal comprehension based on Bormuth's wh-item, a measure of passage independence based on wh-items, the Short Form Test of Academic Aptitude, and the California Achievement Test. The factor analytic results support the conclusion that the MCC measures literal comprehension, a trait that is essentially independent of IQ. However, it was also determined that the MCC had minor loadings on a second, and possibly a third, component related to IQ, inferential reading skills, and language mechanics.

This report presents the results of a series of exploratory factor analyses of a new test of reading comprehension using a multiple-choice cloze format. These analyses are part of a preliminary examination of data gathered on the test in an administration to more than 5,000 students in grades 1-9 in May 1975.

This test development project is concerned with the design and validation of a test of reading comprehension with certain properties that would tend to improve the utility of comprehension testing in the schools. First, instead of providing a fixed test, the intent was to construct a pool of scaled passages and items that could be used to assemble n tests of reading comprehension for a given evaluation purpose with any student or group in grades 1-12. Secondly, the test was to be a measure of literal comprehension or language comprehension per se, as opposed to extant measures of reading comprehension which seem to be psychologically synonymous with higher order reasoning processes (Singer, 1973; Thorndike, 1973-74). Thirdly, the test was to be domain-referenced in the sense that any test assembled from the item and passage pool would represent an unbiased sampling of one or more universes of written discourse. Fourthly, the test was to be based on objective-generative item construction procedures (see Hively, 1974) so that the test construction technology could be economically and reliably reproduced by others.

The standard cloze was initially selected as the item format that offered the most possibilities for building the required test of reading comprehension. Cloze tests are highly passage dependent (the student has virtually no chance of responding correctly unless he reads the passage). The cloze item format offers an objective procedure for the construction of comprehension items--one that can be systematically and widely applied to

-1-

5

samples of written discourse. The item format is also generally coherent
with the ongoing act of reading comprehension if viewed as a constructive
language process (Ryan and Semmel, 1969; Smith, 1975). Since there are no
questions in a cloze item, the test passage remains unaffected by the
idiosyncracies of the item writer. Finally, and of utmost importance for
the construction of a specific measure of comprehension, various deletion
strategies allow for the manipulation of the interaction between reader and
test passage such that the contributions of syntactic, semantic, and reasoning
factors may be controlled.

Although the foregoing features of the standard cloze represent im-
portant advantages, they are considerably offset by problems with validity
and application. On the side of application problems, the standard cloze
format is not readily perceived as a test of reading comprehension, and the
first large-scale attempt to apply the technique as a survey test in the
schools resulted in serious difficulties with interpretation and use of the
data (Hansen and Hesse, 1974). Apparently, the standard cloze is also an
extremely difficult and anxiety-invoking test (Cranney, 1972; Rankin, 1974).
The required length of a cloze passage makes it inconventiently long as a
unit of test assembly. And, the test has the notable disadvantage of
requiring hand scoring.

On the issue of validity, Bormuth (1969) states, "Much of the research
has shown that scores on cloze tests are highly correlated with scores on
standardized tests of reading comprehension ability," but actually reviews
of the literature emphasize low to moderate correlations (Potter, 1968;
O'Reilly et al., 1976). There is also a strong indication in the literature
that the correlation between cloze scores and comprehension scores on
standardized tests is substantially attributable to the concentration of

IQ in both tests (Rankin, 1974).

Some studies of the validity of the cloze as a measure of reading

comprehension indicate that the standard or any-word deletion pattern

unduly weights the syntactic component in a test passage at the expense

of the semantic component (Taylor, 1953; Louthan, 1965; Bickley, Weaver,

and Ford, 1968; and Rankin, 1974). It also appears that responses to the

deletions in a standard cloze test passage are chiefly dependent upon a

surrounding context of 5-10 words, (Taylor, 1956; MacGinitie, 1961),

suggesting insensitivity of the test to the larger ideas that may run

through the passage (Carroll, 1972). Finally, it may be fairly said that

cloze research has generally not been well designed to explore the issue

of the validity of the construct underlying the test, as Ohnmacht, Weaver,

and Kohler (1970) have remarked:

> The fact that responses to cloze tasks reflecting essentially
> gross deletion strategies align themselves with crude mea-
> sures of comprehension does little to shed light upon the
> fundamental nature of comprehension other than to indicate
> that one can measure what passes for comprehension in more
> than one way . . .. Researchers using the cloze procedure
> ought to give careful consideration to language operations
> and to rational operations which are implicit in verbal
> activity and they should construct deletion patterns which
> seem to relate to these operations. Rather than standardizing
> a particular cloze deletion type, exploration of a wider
> range of deletion types which are related to particular
> linguistic and psychological hypotheses is needed. (pp. 215-
> 216)

The present work attempted to meet the exhortations of Ohnmacht et al.

in a rational redesign of the standard cloze item format as a measure of

reading comprehension. This study is apparently the first attempt to

explicitly design a comprehension item to tap the "pure" comprehension of

language factor distinguished by Carroll (1972) from the high level

inferential processes that weigh heavily in standardized tests of reading

comprehension. This comprehension or language factor, referred to here as

7

literal comprehension, is briefly defined as the "apprehension of the grammatical and semantic relations which obtain within and among the sentences of the discourse" (Katz and Fodor, 1963, p. 172). The elaboration of the construct given elsewhere (Schuder, Kidder, and O'Reilly, 1976), assumes that literal comprehension is essentially independent of IQ and is marked by certain types of tests or items included in standardized and other measures of reading comprehension, including some vocabulary measures--particularly those which focus on interpretation of word meanings in context, factual questions about passages, questions about explicit details, questions about implications or entailment relations which hold within a passage, paraphrase questions, and certain types of main idea and title questions.

The Modified Cloze Format: Characteristics and Rationale

The cloze item format designed in this research is technically referred to as an item form (Hively, 1974) that is generally suitable for processing brief, coherent passages into multiple-choice cloze (MCC) items. The MCC passage format, as shown in Figure 1, resembles a standard cloze passage attenuated in length. The passages are generally about 70-80 words long (in grades 1 and 2 they range from 25-45 words), and they exhibit the standard deletion rate of approximately every fifth word. However, inspection of Figure 1 shows that the MCC item form is a lexical cloze: Only nouns, verbs, adjectives, and adverbs have been deleted in contrast with the standard or any-word approach which results in deletion of both structural and lexical items.

-4-

Place Figure 1 about here

Below the test passage appears a set.of 5 response choices for each

deletion, one of which is.the exact word deleted from the passage. All

distractors, it may be noted, would be grammatically plausible in the

position of the deleted word for which they function as distractors.

Distractors for each deletion are generated by a computer program that

randomly accesses sets of words from a 12,400 word vocabulary list within

the constraints of: (a) the grammatical class of the deleted word; (b)

whether the deleted word is a "content" word or is a core (basic) vocabulary

word; and (c) the grade level of the deleted word. Any given distractor

thus functions grammatically but not semantically in the position of the

deleted word, is at the same graded reading level as the passage source,

and is a content specific word or a core word as required to match the

subject matter area to which the word belongs. The core or general vocabulary

lists were compiled from Harris and Jacobson's Basic Elementary Reading

Vocabulary (1972) and the EDL Research and Information Bulletin 5: A

Revised Core Vocabulary (Taylor, Frackenpohl, and White, 1969). The

content-specific word lists were compiled from both the Harris and Jacobson

source and the American Heritage Word Frequency Book (Carroll, Davies, and

Richman, 1971).

The MCC item format preserves many of the advantages of the cloze

technique (e.g., absence of questions and objective item construction) while

potentially enhancing its applicability as a measure of reading comprehension.

Face validity appears to have been considerably enhanced and the 10-item

passage unit is a convenient module for the assembly of a test with 5-10

9

passages of increasing difficulty. The excessive difficulty and ambiguity
of the original cloze testing situation appears to have been considerably
reduced. In fact, the MCC test should generally suffer less from such
sources of invalidity as test anxiety because the test passages do not
function as ordinary test items until the student reaches the point of no-
comprehension with a passage.

The validity of the cloze test has been theoretically improved by
selectively preserving some of the original features of the technique and
substantially modifying others. The every fifth-word deletion pattern has
been maintained because this permits the most thorough and objective
sampling of the ideas and linguistic structures of the test passage without
depriving the student of the information necessary to replace the deleted
words (MacGinitie, 1961; Ramanuskas, 1972). The lexical deletion pattern
should tend to improve validity in several respects. Nouns, verbs, adjec-
tives, and adverbs carry most of the information in a passage, thus focusing
the test on the semantic component and on larger semantic units (Fillenbaum,
1963), without excluding the syntactic component. According to prior
research (Taylor, 1956; Rankin, 1974), the lexical deletion pattern should
also tend to reduce the correlations of the test with IQ while enhancing
correlations with test scores reflecting comprehension of the more "explicit"
meanings of a passage.

The procedure for generating distractors that compete grammatically
but not semantically is specifically designed to limit, insofar as possible,
the context for interpretation of the test passage. Prior experimentation
with the MCC format indicated that the use of semantically interfering
distractors would have the effect of introducing a very difficult voca-
bulary element into the test with the further effect of increasing the

10

correlation of the test with IQ (Cranney, 1972). The distractor design
is also intended to enhance test validity by maintaining the passage
dependency of the test. The use of grammatically equivalent responses
in the MCC item format should function to eliminate the use of grammatical
cues as purely the basis for choosing among distractors. Similarly, the
inclusion of competing content words among certain groups of distractors
should tend to eliminate discontinuities in content as a basis for choosing
among distractors.

Method

The issue of the validity of the literal comprehension construct and
the MCC item format as a measure of the construct was studied in the context
of the annual standardized testing program of a medium-sized urban school
district. This school district contributed 2, 40-minute testing periods
during which the MCC item format, along with an alternate measure of
literal comprehension based on Bormuth's (1970) "wh-item," and a brief
measure of passage independence based on the wh-items were administered.
These measures, together with measures of verbal and non-verbal IQ and
measures of language and reading performance available from the school
district standardized testing program, provided a matrix of test scores
suitable for exploring the construct of literal comprehension via factor
analysis.

The breadth of the test administration, which ranged across several
grade levels, and the variety of the tests available in the study permitted
consideration of several meaningful questions relating to the importance
and properties of the construct of literal comprehension. Chief among these
was the question of whether factor analysis would verify a substantial
literal comprehension factor that was generalizable across a large number

11

-7-

of MCC test passages and that was similarly constituted across several grade levels of the study population. In addition it was expected that this literal comprehension factor would be essentially independent of the IQ and passage independence measures available in the study and would be marked by substantial loadings on other reading tests that appear to measure a similar factor or are related to the factor.

Sample

The original study sample consisted of 5,722 students in grades 1-9, with roughly 500-750 students in each grade level. Students were grouped into subsamples for the analysis based on the test levels in the California Achievement Test (CAT) battery. The present study is based on the response data available for the first three subsamples in grades 1-6 where IQ scores were available. Subsample I consisted of 456 first-graders, subsample II had 972 second and third graders, and subsample III had 1697 students in grades 4-6.

Instruments

The test scores available for the analyses for each subsample are listed in clusters in Table 1 under each CAT level, along with descriptive data for each score. Test scores that are expected to strongly mark the literal comprehension factor are underlined. A brief description of each test score cluster follows.

- -

Place Table 1 About Here

- -

12

-8-

Multiple-Choice Cloze Test. The MCC test forms consisted of two sets of 12 parallel test forms, one set for grades 1-3 and one set for grades 4-6. The test forms were systematically assembled from a pool of 353 MCC passages drawn from basal readers and literature texts for grades 1-10. Each set of forms was assigned a range of readability levels in the cloze passage pool (passages in the pool are ordered on readability) and within these ranges of readability consecutive pairs of readability levels were regarded as sampling units (except at grades 1-3 where the first two readability levels were treated as separate sampling units). Test forms for a given grade level range were then constructed by sampling without replacement 6 ordered passages, one from each consecutive sampling unit. The order of passage readability was maintained in the test form. Each test form contained 39, 41, or 60 items presented to the testee in groups of 3, 5, or 10 items. The shorter test forms were at grades 1-3 where the first 3 passages in a form were 25-45 words long with 3 or 5 items per passage.

The MCC test yields 4 subscores corresponding to the grammatical classes of the words deleted in a cloze passage. Due to the distribution of grammatical classes in the passages, the noun score has the largest mean and variance, followed by the verb score and the scores for adjectives and adverbs. Internal reliabilities (KR-20) for the MCC test forms in grades 1-3 and 4-6 ranged from .94-.97, with a median reliability of .96.

The Wh-Item Test. Because the standardized measure of reading used in the study was, in many respects, an ambiguous criterion for the MCC as a measure of literal comprehension, an alternative measure of the construct was developed. Forms of this test, called the Wh-Item Test were assembled from a pool of some 300 ordered passages and 3,000 main idea and wh-items

13

using a design for the selection of test passages that was virtually
identical to that used for the assembly of the cloze test forms. Five wh-
items were selected from the 8 wh-items available for each of the 6
ordered passages in a test form so that there were equal distributions of
wh-item types across test forms. The wh-item types are: how, what (noun),
what (verb), when, where, which, who, why. This procedure resulted in two
sets of uniform, 30-item tests in each test level that paralleled the
Multiple-Choice Cloze tests in number of passages and range of passage
difficulty.

The Wh-Item Test yields 8 subscores corresponding to the wh-item types
represented in the test. Internal reliabilities (KR-20) for the Wh-Item
Test in grades 1-3 ranged from .90-.94, with a median reliability of .91;
and in grades 4-6 from .85-.94, with a median reliability of .93.

Test Wiseness Test. Because the MCC and Wh-Item Test forms were
considered to be passage-dependent measures of reading comprehension, a
special test was constructed to test this assumption. The design of this
test, referred to as the Test-Wiseness Test, paralleled the MCC and Wh-Item
Test form designs. The questions, not the passages, in each Wh-item Test
form were pooled separately for grades 1-3 and 4-6. A set of 12 test forms
was then constructed for each grade level range by systematically drawing
12 items from this pool for each test form. Care was also taken to represent
the related passage difficulties for the items and the 8 types of wh-items
in a test form in an attempt to create parallel tests. The relationship
between scores on this Test-Wiseness measure and scores on the Wh-Item
Test provides some indication of the extent to which student's responses
on the latter test are dependent on reading the associated test passages.
This test also provides some indication of the extent to which this form

14

of test-wiseness affects responses on the Multiple-Choice Cloze test.

The Test-Wiseness Test yields a single score. Internal reliabilities for the Test-Wiseness Test in grades 1-3 ranged from .13-.79, with a median reliability of .68; and in grades 4-6 from .29-.76, with a median reliability of .70.

Short Form Test of Academic Aptitude. The Short Form Test of Academic Aptitude (SFTAA) is a group-administered intelligence test that yields language and non-language IQ's. This test, administered by the school district to students in grades 1-6, along with the California Achievement Test, permitted study of the relationship between IQ and the literal comprehension test across the study subsample.

California Achievement Test. The various CAT reading and language test scores used in the study were previously listed in Table 1 by CAT test level. These subscores, rather than the lengthier and more reliable CAT skill scores (major headings in Table 1) were used in order to provide a less ambiguous basis for marking the expected literal comprehension factor—as opposed to an inferential factor which might be expected to appear, marked by IQ and such CAT subtests as generalizations and inferences. However, preliminary correlational analysis and inspection of the CAT comprehension items indicated that this approach did not satisfactorily resolve the CAT comprehension section into independent literal and non-literal subtests. The CAT comprehension section at every level appeared to be overall considerably more "literal" than was expected, in light of the subtest labels. The supposedly inferential subtests were substantially contaminated with literal items and vise versa. Both "literal" and "non-literal" subtests also contained numbers of items that appeared to be passage independent (could be answered without reading the passage).

15

Consequently, in the hopes of further disambiguating the criteria for marking literal comprehension in the factor analysis, the CAT comprehension items were re-classified into three new subtest scores: (a) items that appeared to measure literal comprehension and were passage dependent; (b) items that appeared to measure literal comprehension but were passage independent; and (c) non-literal items or items that seemed to reflect higher order, inferential processes. The CAT subtests based on the literal-non-literal and passage-dependent item classification are identified as the CAT Item Classification cluster in Table 1.

Analysis

The data set available on the foregoing test scores was organized separately for analysis by each CAT level identified in Table 1. To permit analysis across the test forms constructed for the study at each level, the raw scores for the Wh-Item, MCC and, Test-Wiseness Test forms were converted to z scores based on the score distribution for each test form in a test level. Subsequently, negative values were eliminated by applying a linear transformation to each set of obtained z scores. The resultant scores from any of the foregoing tests in a test level were thereafter treated as having come from equivalent test forms and were combined as required for the analyses by CAT level.[1]

_____

[1] This approach to test equating, though somewhat unorthodox, is defensible on several grounds. The general shapes, means, and the standard deviations of the distributions of the Wh-Item and MCC test scores were very similar from form to form in a test level (usually the average raw score difference from form to form was less than one-fourth of a standard deviation), the internal reliabilities of each form were consistently high, and the tests had been systematically assembled to be parallel in order and range of readability level. This approach, however, was less defensible for the Test-Wiseness Test which varied from form to form in reliability and in the distribution of scores.

16

The factor analyses were organized in three stages: (1) the first

stage factor analyzed only the MCC and wh-item subscores; (2) the second

stage then added the conventional CAT subscores and the IQ scores identified

in Table 1 to the analysis; and (3) the third stage replaced the CAT

comprehension scores with the CAT Item Classification scores and added the

Test-Wiseness score. In each stage of the analysis, the various test scores

were intercorrelated by subsample and the resulting matrices subjected to

principal components analysis with ones in the diagonals. Components with

eigenvalues > 1.00 were then rotated to the varimax criterion. These

analyses were then rerun with only the noun and verb scores used to represent

the MCC test in the hope of lending further clarification to the results.

To evaluate the expectation that the factors in the analyses would be

correlated, the factor analyses were run again and rotated to the oblique

criterion. The resulting correlations between the obtained factors are

of theoretical interest here, but the factor matrices are not reported

because the oblique results were nearly identical to the orthogonal findings.

<div align="center">Results</div>

Stage 1

The results of the factor analyses at stage 1 are given in order by

subsample in Tables 2-4. The rotated factor matrices in the first analysis

indicate a consistent tendency for the MCC test to split into two factors

across subsamples: (a) I is the more important factor and is defined

primarily by the noun and verb scores and the Wh-Item subscores; and (b)

II is defined most strongly by the adjective and adverb subscores in grade 1

and very strongly by all MCC subscores in the other subsamples. Deleting the

adjective and adverb subscores, as shown in the second set of factor matrices,

reduces the matrix to one factor and sharply increases the contribution of the

<div align="center">17</div>

cloze noun and verb subscores to I. This analysis was originally designed to reveal that II was attributable to the adjective and adverb subscores, but actually the pattern of results indicates an increasing contribution of all four MCC subscores to this second cloze component across subsamples.

_____

Place Tables 2-4 About Here

_____

Stage 2

The results for the stage 2 factor analyses are shown in Tables 5-7 by subsample. Turning to the first factor matrix for grade 1 in Table 5, three factors obtain, the first of which is identifiable as literal compre- hension, being marked by the CAT Words in Context subtest, the cloze noun score, the CAT comprehension subtests, the cloze verb score, the various Wh-Item subscores, and the CAT subtests for Picture-Word Association, Language Mechanics and Language Usage. The last three subtests are not of particular importance in defining factor I. As expected, the IQ subscores load at very low levels with I.

_____

Place Tables 5-7 About Here

_____

The second factor is composed largely of the CAT phonological, ortho- graphic and word recognition skills, with moderate loadings on the factor for the IQ subscores and the CAT language test scores. This factor seems to reflect a combination of the pre-reading skills and general verbal ability that are important components in learning to read.

18

The third factor appears to be the second component of the MCC identified in stage 1 which can be seen here to be independent of IQ. The fourth factor is primarily defined by the CAT Sentence-Picture Association subscore.

Dropping the adjective and adverb subscores in the second factor matrix increases the loading of the cloze noun and verb scores in I and eliminates the second component of the cloze as in Stage I. In the oblique solution, the correlations among the factors were generally low (RI • II = .41 and RI • III = .38), supporting the hypothesis that the literal comprehension factor would be essentially independent of inferential processes.

The pattern of results for Level I is considered to be generally supportive of theoretical expectations although there is the apparent inconsistency of the loading of the CAT Inferences subtest on I and the failure of a CAT "inferences" factor to appear in the matrix. These inconsistencies seem to be resolved by the fact that inferential processes are represented only weakly in the Level I CAT comprehension section, there being only 8 items thus classified. Moreover, in the process of completing the CAT Item Classification scores, many of these items were seen as doubtful measures of inference.

The rotated factor matrices for grades 2 and 3 are shown in Table 6. The first factor seems to be clearly a literal comprehension factor with moderate to high loadings on the cloze and Wh-Item tests, and the CAT Word Recognition, Words in Context, and Facts subtests. The IQ subtests are virtually uncorrelated with Factor I and the inferential subtests have moderate loadings with the factor. Factor II, which has high loadings for all four cloze subscores, appears to be a complex of variables involving language skills, IQ, and virtually all of the vocabulary and comprehension

19

-15-

sections of the CAT. This factor appears to combine the more than "strictly literal" component of the cloze, that is thought to be reflected in the tendency of the MCC test to split into a second component, with the even higher order reasoning processes reflected in the IQ subscores.

As before, dropping the adjective and adverb component from the test score matrix in the second analysis raises the contribution of the cloze to I, but does not otherwise change the interpretation of the results. The failure of the second cloze component to be resolved as a factor independent of IQ appears to explain in part the substantial correlation between factors I and II (RI • II = .64) in the oblique solution.

The factor matrices for the grade 4-6 subsample are shown in Table 7, and as will be seen, these results are more consistent with expectations. As before, I is clearly interpretable as a literal comprehension factor, but here the loadings of the MCC with I are generally higher than in previous levels of the analysis. Factors II and III appear to have resolved the conglomeration of language and reasoning skills in factor II of the previous analysis into two separate factors, each of which has a minor cloze contribution. Factor II is primarily a language factor, while III is largely an IQ or reasoning factor marked by moderate loadings for the CAT vocabulary and Comprehension subtests.

Dropping the adjective and adverb score in the second analysis has no appreciable effect on the pattern of results. The correlations among the factors from the oblique solution were somewhat lower than in the previous analysis (RI • II = .42; RI • III = .54).

20

Stage 3

Since the results for stage 3, shown in Tables 8-10, closely parallel the findings of the previous level of analysis, they are discussed here as a group, with a focus on the possible contribution of the CAT Item Classification to theoretical clarity. In the grade 1 subsample, there is a tendency for the literal comprehension subscore to load on factor I more substantially than the non-literal subscore. A similar relationship is found in the grade 2 and 3 results, but the literal comprehension score also loads about equally on factors I and II. In the grade 4-6 subsample, the pattern is somewhat more consistent with expectations in that the non-literal comprehension subscore loads at a low level with factor I and at a moderate level with factor III--the IQ or reasoning factor. However, the literal comprehension subscores load about equally with factors I and III. The Test-Wiseness score added to this stage of the analysis fails to relate substantially to any of the factors.

Place Tables 8-10 About Here

Discussion

In retrospect, the present study represents a highly complex background against which a tentative and still vague conceptualization of literal comprehension was explored. The analysis of factor structures across different age-graded samples and variable test criteria constituted a complex interacting context involving developing cognitive abilities; shifts in the psycholinguistic meaning of the test criteria used, resulting very likely in changes in the types of skills tapped; and changes within and between subsamples in the demands made by the MCC and Wh-Item formats on students' syntactic and semantic competence. It is unreasonable to expect

21

-17-

any clean set of results given this context for exploration and, certainly,
a somewhat mixed set of results ensued. However, the results seem to be
sufficiently consistent to conclude that the conceptualization for the study
is in the right direction and to further offer a few tentative generalizations.

The data appear to support the conclusion that the MCC format is in
part a measure of a restricted form of reading comprehension that is
essentially independent of IQ. This form of comprehension appears to be
interpretable as the apprehension of the "strictly literal" meanings con-
tained in sentences and phrases as measured by reading tests that focus on
factual questions, questions about explicit details, and questions about
interpretation of meanings within the context of isolated sentences and
phrases. More tenuously related to the data is the conclusion that the
MCC format is composed of a second and possibly a third component that
reflects other than "strictly literal" comprehension processes. It was
apparently too much to expect that appropriate criteria elucidating this
second component of the cloze would be found in the CAT, even with an
arduous re-classification of the comprehension items in the test.

The next stage of research on the MCC format must obviously be concerned
with the development of a broader range of test or performance criteria
specifically designed to tap the more expansive implications of the MCC
format as a broad and generalizable measure of literal comprehension.
Recent progress in clarifying the construct of literal comprehension in
Schuder et al. (1976), beyond the admittedly crude conceptualization that
guided the re-classification of CAT comprehension items used here, provides
a number of important leads for constructing these test criteria. In
addition, research along these lines must be concerned with measuring the
syntactic and semantic demands made by the test passages on the testee.

22

Clearly, the complexity of the test passages in a cloze test will influence the correlations between the test and other reading and cognitive performance criteria.

The findings presented here further demonstrate the futility, from a theoretical point of view, of correlating cloze test scores with overall scores from standardized measures of reading comprehension or with similar home-grown measures. Judging by the CAT, such tests are a complex collection of item types, whose psycholingustic nature is not especially revealed by the test maker. That the psycholinguistic ambiguity of the CAT may generalize, at least in part, to the products of other major reading test makers was shown in a recent study by Tuinman (1973-74). Tuinman's analysis showed that several well-known standardized comprehension tests had substantial numbers of items that were not passage dependent—a problem that was clearly apparent in the detailed examination of the CAT in the present study.

Finally, it seems that the present study provides tentative support for Carroll's (1972) contention that language comprehension, or literal comprehension as it is called here, could be isolated from inferential or reasoning processes by more careful test construction procedures. Identified as a factor(s) in the present study, literal comprehension accounted for well more than a minuscule proportion of the variance of the various tests used in the factor analyses. In addition, the contention that literal comprehension would be essentially independent of IQ was largely supported by the data, particularly if the rationale concerning the literal versus non-literal content of the CAT is accepted. Other studies have found much higher correlations among factors that are presumed to make up tests of comprehension and reasoning (Bormuth, 1969; Davis, 1968; Spearritt, 1972)--

23

-19-

so much so that reading comprehension and reasoning have been equated
(Thorndike, 1973-74).

Carroll's hypothesis is thus very much worthy of further investigation,
particularly since the bifurcation of reading skills into two basic processes
has broad implications for reading instruction.

24

REFERENCES

Bickley, A., Weaver, W., & Ford, F. Information removed from multiple-
  choice item responses by selected grammatical categories. Psychological
  Reports, 1968, 23, 613-614.

Bormuth, J. Factor validity of cloze tests as measures of reading compre-
  hension ability. Reading Research Quarterly, 1969, 4, 358-365.

Bormuth, J. On the theory of achievement test items. With an appendix by
  P. Menzel: On the linguistic bases of the theory of writing items.
  Chicago: University of Chicago Press, 1970.

Carroll, J. Defining language comprehension: Some speculations. In
  R. Freedle & J. Carroll (Eds.), Language comprehension and the acquisition
  of language. New York: Wiley, 1972.

Carroll, J., Davies, P., & Richman, B. The American Heritage word frequency
  book. New York: Houghton-Mifflin, 1971.

Cranney, A. The construction of two types of cloze reading tests for college
  students. Journal of Reading Behavior, 1972, 5, 60-64.

Davis F. Research in comprehension in reading. Reading Research Quarterly,
  1968, 3, 499-545.

Fillenbaum, S., Jones, L., & Rapoport, A. The predictability of words and
  their grammatical classes as a function of rate of deletion from a
  speech transcript. Journal of Verbal Learning and Verbal Behavior,
  1963, 2, 186-194.

Hansen, L.H., & Hesse, K.D. A pilot reading literacy assessment of Madison
  Public School students: Final report. Madison, Wisconsin: The
  Madison, Wisconsin Public Schools, The Department of Research and
  Development, 1974.

Harris, A., & Jacobson, M. Basic elementary reading vocabularies. New
  York: Macmillan, 1972.

Hively, W. Domain-referenced testing: Part one, basic ideas, introduction
  to domain-referenced testing. Educational Technology, 1974, 14 (6), 5-10.

Katz, J. & Fodor, J. The structure of a semantic theory. Language, 1963,
  39, 170-210.

Louthan, V. Some systematic grammatical deletions and their effects on
  reading comprehension. English Journal, 1965, 54, 295-299.

MacGinitie, W. Contextual constraint in English prose paragraphs. The
  Journal of Psychology, 1961, 51, 121-130.

Ohnmacht, F., Weaver, W., & Kohler, E. Cloze and closure: A factorial
  study. The Journal of Psychology, 1970, 74, 205-217.

25

O'Reilly, R.P., Kidder, S.J., & Schuder, R.T. A Proposal for Validation and Refinement of Measures of Literal Comprehension in Reading for Use in Policy Research and Classroom Management (Photocopy). Albany, N.Y.: The University of the State of New York, The State Education Department, Division of Research, 1975.

Potter, T. A taxonomy of cloze research, part I: Readability and reading comprehension (Tech. Rep. No. 11). Los Angeles: Southwest Regional Laboratory for Educational Research and Development, 1968.

Ramanauskas, S. Contextual constraints beyond a sentence on cloze responses of mentally retarded children. American Journal of Mental Deficiency, 1972, 77, 338-345.

Rankin, E. The cloze procedure revisited. In P.L. Nacke (Ed.), Interaction: Research and practice for college-adult reading, Twenty-third Yearbook of the National Reading Conference, 1974.

Ryan, E., & Semmel, M. Reading as a constructive language process. Reading Research Quarterly, 1969, 5, 59-83.

Schuder, R.T., Kidder, S.J., & O'Reilly, R.P. Defining and measuring the literal comprehension of written discourse. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1976.

Singer, H. IQ is and is not related to reading. Paper presented at the Annual Convention of the International Reading Association, Denver, Colorado, May 6, 1973.

Smith, F. Comprehension and learning: A conceptual framework for teachers. New York: Holt, Rinehart, & Winston, 1975.

Spearritt, D. Identification of subskills of reading comprehension by maximum likelihood factor analysis. Reading Research Quarterly, 1972, 8, 92-111.

Taylor, S., Frackenpohl, H., & White, C. A revised core vocabulary. A basic vocabulary for grades 1-8, an advanced vocabulary for grades 9-13 (Research and Information Bulletin No. 5). Huntington, N.Y.: Educational Developmental Laboratories, March 1969.

Taylor, W. "Cloze procedure": A new tool for measuring readability. Journalism Quarterly, 1953, 30, 414-438.

Taylor, W. Recent developments in the use of the cloze procedure. Journalism Quarterly, 1956, 33, 42-48.

Thorndike, R. Reading as reasoning. Reading Research Quarterly, 1973-1974, 9, 135-147.

Tuinman, J. Determining the passage dependency of comprehension questions in 5 major tests. Reading Research Quarterly, 1973-1974, 9, 206-223.

26

Figure 1

THE YOUNG WHALE

The young whale tapped his teeth and _____ Coos Bay. He had been _____ in January, a magnificent _____ of sixteen feet. Upon his _____ in the whale world, he had been _____ nuzzled by his giant _____, who, without arms or _____ with which to hug him, _____ her love by circling him. She _____ him to the surface to _____, then, tipping her body, she showed him where he would find her milk.

○ a. circled
b. loaned
c. obeyed
d. became
e. famed

○ a. fern
b. lap
c. puppet
d. beech
e. mother

○ a. thankful
b. nervous
c. slow
d. foul
e. born

○ a. sauces
b. feet
c. cuts
d. hills
e. inns

○ a. hawk
b. quail
c. pipe
d. male
e. flea

○ a. computed
b. decorated
c. copied
d. expressed
e. repaired

○ a. scorn
b. location
c. raccoon
d. blister
e. arrival

○ a. stitched
b. married
c. glued
d. led
e. lit

○ a. indignantly
b. immediately
c. warily
d. hoarsely
e. viciously

○ a. ache
b. bow
c. blow
d. add
e. fade

27

Table 1

Means, Standard Deviations, and Number of Test Items for Tests Included in the Factor Analyses

Test	CAT Level I (N = 456)			CAT Level II (N = 972)			CAT Level III (N = 1697)		
	Mean	S.D.	Items	Mean	S.D.	Items	Mean	S.D.	Items
Multiple-Choice Cloze Test[a] (MCC)	49.58	9.97	41[b]	49.58	9.89	41	49.80	9.82	60
Wh-Item Test	49.59	9.94	30	49.68	9.83	30	49.71	9.90	30
Test-Wiseness Test	49.71	9.79	12	49.53	9.86	12	49.61	9.91	12
Short Form Test of Academic Aptitude									
Language IQ	101.87	13.90	45	100.12	15.94	45	97.20	18.60	45
Non-language IQ	104.08	16.38	40	103.72	16.30	40	102.90	19.07	40
California Achievement Test									
Reading Vocabulary – ADSS[c]	316.11	34.52	92	369.38	49.65	40	429.43	69.41	40
Sentence-Picture Association	9.87	.45	10						
Beginning Sounds	8.02	1.82	10						
Ending Sounds	8.96	1.43	10						
Letter Recognition	14.54	1.57	15						
Word Forms	8.41	2.09	10						
Picture-Word Association	7.19	2.42	10						
Word Recognition	10.04	2.47	12	18.50	2.49	20	24.83	8.82	40
Words in Context	5.45	3.90	15	14.42	4.84	20			
Reading Comprehension-ADSS	297.08	52.07	24	388.20	67.32	45	460.03	72.15	43
Facts	3.29	2.06	7	8.61	3.34	12	8.52	2.31	11
Interpretation	2.12	1.54	6	5.69	2.62	9	9.76	3.49	15
Relationships							1.52	1.16	4
Generalizations				3.87	1.82	6	3.97	1.94	9
Inferences	2.29	1.74	8	5.19	2.34	8	1.85	1.13	4

Table 1 (Continued)

Test	CAT Level I (N - 456)			CAT Level II (N = 972)			CAT Level III (N = 1697)		
	Mean	S.D.	Items	Mean	S.D.	Items	Mean	S.D.	Items
Comprehension Item Classifications									
Literal Comp./Passage Depend.	4.17	2.70	11	11.94	4.49	17	4.49	1.86	8
Literal Comp./Passage Indep.				4.41	2.25	7	5.32	2.07	9
Non-literal Comprehension	1.84	1.41	6	5.71	2.61	9	3.04	1.80	8
Language Mechanics-ADSS	303.50	58.42	37	385.05	72.95	66	464.43	98.72	80
Language Usage- ADSS	324.93	63.11	20	391.19	73.41	25	454.08	88.92	41
Standard English	11.70	3.88	20	14.70	5.18	25	14.42	3.52	24
Sentence Structure							3.35	1.22	5
Sentence Parts and Funct.							1.76	1.65	7
Transformations							2.57	1.29	5

[a]Scores on the MCC, Wh-Item, and Test-Wiseness Tests were standardized separately within CAT Levels I, II, and III to have a mean of 50 and standard deviation of 10. Four subscores (Noun, Verb, Adjective, and Adverb) were included in the MCC Test with nouns and verbs containing the greater proportion of items. The Wh-Item Test contained eight subscores (How, What-N, What-V, When, Where, Which, Who, and Why), each averaging between three and four items per test form.

[b]Two of the twelve forms in CAT Level I and -II contained 39 items.

[c]Achievement Development Scale Scores (ADSS) were derived by CTB/McGraw-Hill from a single equal-interval score scale across all grades for use with all levels and forms of the CAT.

Table 2

Factor Loadings for Multiple-Choice Cloze and Wh-Item Subscores: Grade 1
(N = 456)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores			Unrotated matrix including Multiple-Choice Cloze noun and verb subscores only	
	I	II	Final Communality	I	Final Communality
Multiple-Choice Cloze:					
Noun	.57	.61	.70	.77	.59
Verb	.50	.63	.64	.71	.51
Adjective		.77	.61		
Adverb		.73	.54		
Wh-Item:					
How	.64		.47	.68	.47
What (N)	.67		.49	.69	.48
What (V)	.69		.49	.68	.46
When	.55		.32	.56	.31
Where	.65		.44	.64	.41
Which	.70		.53	.72	.52
Who	.75		.56	.72	.51
Why	.58		.35	.58	.33
Eigenvalue	4.88	1.25		4.59	
Cumulative Proportion of Total Variance	.41	.51		.46	

30

Table 3

Factor Loadings for MCC and Wh-Item Subscores: Grades 2 and 3

(N = 972)

Subscores	Varimax matrix including all four Multiple Choice Cloze subscores			Unrotated matrix including Multiple Choice Cloze noun and verb subscores only	
	I	II	Final Communality	I	Final Communality
Multiple Choice Clozes:					
Noun	.50	.76	.82	.82	.67
Verb	.50	.74	.80	.82	.67
Adjective		.76	.66		
Adverb		.75	.58		
Wh-Item:					
How	.73		.56	.72	.52
What (N)	.68		.53	.72	.52
What (V)	.67	.30	.54	.73	.53
When	.56		.39	.63	.39
Where	.68		.52	.71	.50
Which	.74		.60	.75	.56
Who	.75		.61	.76	.58
Why	.63		.48	.69	.48
Eigenvalue	6.07	1.03		5.42	
Cumulative Proportion of Total Variance	.51	.59		.54	

31

Table 4

Factor Loadings for Multiple-Choice Cloze and Wh-Item Subscores: Grades 4, 5, and 6
(N = 1697)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores			Unrotated matrix including Multiple-Choice Cloze noun and verb subscores only	
	I	II	Final Communality	I	Final Communality
Multiple-Choice Cloze:					
Noun	.42	.83	.86	.82	.68
Verb	.41	.82	.85	.81	.66
Adjective	.34	.82	.78		
Adverb		.79	.67		
Wh-Items:					
How	.66	.30	.52	.71	.51
What (N)	.66		.50	.68	.47
What (V)	.63		.47	.68	.46
When	.68		.51	.68	.46
Where	.66		.51	.69	.48
Which	.72		.59	.74	.55
Who	.65		.51	.70	.50
Why	.64	.33	.52	.72	.52
Eigenvalue	6.26	1.03		5.28	
Cumulative Proportion of Total Variance	.52	.61		.53	

Table 5

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Short Form Test of Academic Aptitude, and California Achievement Test Subscores: Grade 1

(N = 456)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores					Varimax matrix including Multiple-Choice Cloze noun and verb subscores only			
	I	II	III	IV	Communality	I	II	III	Communality
Multiple-Choice Cloze:									
Noun	-.58		-.54		.69	.73			.59
Verb	-.48		-.55		.61	.64			.48
Adjective			-.75		.61				
Adverb			-.71		.51				
Wh-Item:									
How	-.67				.48	.65			.43
What (N)	-.64				.49	.63			.47
What (V)	-.65				.47	.63			.43
When	-.54				.34	.53			.32
Where	-.65				.44	.62			.40
Which	-.66				.49	.66			.48
Who	-.69				.54	.65			.49
Why	-.55				.32	.52			.30
Short Form Test of Academic Aptitude:									
Language IQ		.48		.44	.51		-.47	.45	.50
Non-Language IQ	-.36	.51		.30	.49	.37	-.50	.32	.49
California Achievement Test:									
Sentence-Picture Association				.84	.72			.83	.70
Beginning Sounds		.78			.70	.30	-.77		.69
Ending Sounds		.73			.58		-.73		.56
Letter Recognition		.63			.44		-.64		.44
Word Forms		.67			.51		-.67		.51
Picture-Word Association	-.45	.70			.71	.48	-.69		.71
Word Recognition		.79			.68		-.79		.68
Words in Context	-.70	.31	-.31		.69	.77			.70
Facts	-.64		-.35		.61	.72			.60
Interpretation	-.66				.54	.70			.54
Inferences	-.63				.50	.69			.51
Language Mechanics	-.42	.52			.57	.45	-.51		.57
Standard English	-.39	.65			.65	.39	-.64	.33	.64
Eigenvalue	10.08	2.42	1.26	1.05		9.85	2.34	1.05	
Cumulative Proportion of Total Variance									
Variance	.37	.46	.51	.55		.39	.49	.53	

33

Table 6

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Short Form Test
of Academic Aptitude and California Achievement Test Subscores: Grades 2 and 3
(N = 972)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores			Varimax matrix including Multiple-Choice Cloze noun and verb subscores only		
	I	II	Communality	I	II	Communality
Multiple-Choice Cloze:						
Noun	.54	.68	.76	.61	.60	.72
Verb	.54	.66	.72	.60	.58	.70
Adjective	.35	.61	.49			
Adverb		.49	.29			
Wh-Item:						
How	.70		.53	.70		.52
What (N)	.65	.30	.51	.65		.50
What (V)	.67		.52	.69		.52
When	.55		.37	.56		.37
Where	.66		.50	.67		.49
Which	.73		.58	.73		.57
Who	.73		.60	.73		.59
Why	.60	.31	.46	.61		.46
Short Form Test of Academic Aptitude:						
Language IQ		.73	.56		.78	.63
Non-language IQ		.72	.53		.79	.62
California Achievement Test:						
Word Recognition	.61	.36	.49	.61	.35	.49
Words in Context	.68	.57	.78	.70	.54	.78
Facts	.61	.64	.78	.65	.61	.79
Interpretation	.53	.67	.73	.56	.64	.73
Generalizations	.49	.68	.70	.53	.65	.71
Inferences	.55	.64	.72	.59	.61	.72
Language Mechanics	.42	.66	.61	.45	.65	.62
Standard English	.55	.64	.71	.58	.62	.72
Eigenvalue	11.72	1.24		11.07	1.21	
Cumulative Proportion of Total Variance	.53	.59		.55	.61	

34

Table 7

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Short Form Test
of Academic Aptitude, and California Achievement Test Subscores: Grades 4, 5, and 6
(N = 1697)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores				Varimax matrix including Multiple-Choice Cloze noun and verb subscores only			
	I	II	III	Communality	I	II	III	Communality
Multiple-Choice Cloze:								
Noun	.71	.35	.33	.73	.66	.34	.37	.69
Verb	.70	.35	.32	.72	.65	.34	.37	.68
Adjective	.64	.33		.61				
Adverb	.56			.44				
Wh-Item:								
How	.70			.52	.70			.53
What (N)	.64			.46	.65			.48
What (V)	.61			.44	.62			.46
When	.65			.45	.66			.47
Where	.67			.48	.68			.50
Which	.71			.54	.72			.57
Who	.68			.49	.69			.51
Why	.67			.50	.67			.51
Short Form Test of Academic Aptitude:								
Language IQ			.82	.76			.83	.75
Non-Language IQ			.81	.72			.81	.71
California Achievement Test:								
Words in Context	.47	.42	.59	.75	.46	.42	.60	.75
Facts	.56	.30	.49	.64	.55	.30	.50	.64
Interpretation	.51	.42	.52	.71	.50	.42	.53	.71
Relationships		.32	.47	.40		.33	.47	.40
Generalizations		.40	.56	.56		.41	.56	.56
Inferences			.54	.39		.30	.53	.38
Language Mechanics	.36	.65	.35	.68	.35	.66	.36	.69
Standard English		.62		.50		.60		.49
Sentence Structure		.67		.49		.68		.50
Sentence Parts and Functions		.64		.45		.64		.46
Transformations		.65		.44		.66		.44
Eigenvalue	10.82	1.84	1.18		9.86	1.83	1.17	
Cumulative Proportion of Total Variance	.43	.51	.55		.43	.51	.56	

35

Table 8

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Test Wiseness Short Form Test of Academic Aptitude and California Achievement Test Subscores Including Comprehension Item Classifications: Grade 1
(N = 456)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores					Varimax matrix including Multiple-Choice Cloze noun and verb subscores only				
	I	II	III	IV	Communality	I	II	III	IV	Communality
Multiple-Choice Cloze:										
Noun	.55		.55		.68	.71				.59
Verb	.45	.30	.57		.62	.62				.51
Adjective			.75		.60					
Adverb			.69		.49					
Wh-Item:										
How	.67				.49	.70				.52
What (N)	.64				.49	.66				.51
What (V)	.65				.47	.65				.46
When	.53				.33	.51				.32
Where	.65				.45	.62				.41
Which	.66				.50	.68				.50
Who	.69				.55	.67				.51
Why	.54				.32	.53				.31
Test Wiseness			.34		.16				.87	.78
Short Form Test of Academic Aptitude:										
Language IQ		.48		.44	.50		.46	.46		.50
Non-Language IQ	.35	.52			.49	.39	.49	.33		.51
California Achievement Test:										
Sentence-Picture Association				.84	.72			.84		.71
Beginning Sounds		.78			.70		.78			.70
Ending Sounds		.73			.57		.73			.58
Letter Recognition		.62			.43		.64			.44
Word Forms		.67			.50		.67			.51
Picture-Word Association	.44	.70			.71	.48	.69			.71
Word Recognition		.79			.69		.78			.69
Words in Context	.67	.32	.35		.69	.74	.30			.69
Language Mechanics	.41	.53			.56	.46	.50	.33		.57
Standard English	.39	.66			.65	.40	.63			.65
Comprehension Item Classification										
Lit. Comp./Pass. Dependent	.68		.39		.69	.75				.69
Non-Literal Comp.	.53		.33		.43	.57	.08			.49
Eigenvalue	9.71	2.41	1.27	1.05		9.48	2.31	1.05	1.01	
Cumulative Proportion of Total Variance	.36	.45	.50	.53		.38	.47	.51	.55	

Table 9

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Test Wiseness,
Short Form Test of Academic Aptitude and California Achievement Test Subscores
Including Comprehension Item Classifications: Grades 2 and 3
(N = 972)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores			Varimax matrix including Multiple-Choice Cloze noun and verb subscores only		
	I	II	Communality	I	II	Communality
Multiple-Choice Cloze:						
Noun	.54	.68	.76	.61	.59	.72
Verb	.54	.66	.73	.61	.58	.70
Adjective	.35	.61	.49			
Adverb		.50	.30			
Wh-Item:						
How	.70		.54	.70		.52
What (N)	.65	.30	.51	.65		.50
What (V)	.67		.52	.69		.52
When	.55		.37	.57		.37
Where	.67		.50	.67		.50
Which	.73		.58	.73		.57
Who	.73		.60	.73		.59
Why	.60	.31	.46	.61		.46
Test Wiseness			.09			.09
Short Form Test of Academic Aptitude:						
Language IQ		.73	.56		.79	.65
Non-Language IQ		.72	.53		.79	.64
California Achievement Test:						
Word Recognition	.60	.36	.49	.61	.35	.49
Words in Context	.68	.56	.78	.70	.53	.78
Language Mechanics	.43	.66	.62	.46	.64	.62
Standard English	.55	.63	.71	.59	.61	.72
Comprehension Item Classification						
Lit. Comp./Pass. Depend.	.61	.64	.79	.65	.60	.79
Lit. Comp./Pass. Indep.	.46	.68	.67	.50	.64	.65
Non-Literal Comprehension	.56	.66	.75	.60	.63	.75
Eigenvalue	11.12	1.24		10.46	1.20	
Cumulative Proportion of Total Variance	.51	.56		.52	.58	

37

Table 10

Factor Loadings for Multiple-Choice Cloze, Wh-Item, Test Wiseness,
Short Form Test of Academic Aptitude and California Achievement Test Subscores
Including Comprehension Item Classifications: Grades 4, 5, and 6
(N = 1697)

Subscores	Varimax matrix including all four Multiple-Choice Cloze subscores				Varimax matrix including Multiple-Choice Cloze noun and verb subscores only			
	I	II	III	Communality	I	II	III	Communality
Multiple-Choice Cloze:								
Noun	.69	.37	.33	.73	.65	.36	.36	.68
Verb	.68	.37	.33	.71	.64	.36	.37	.67
Adjective	.63	.35	.30	.61				
Adverb	.55			.44				
Wh-Item:								
How	.70			.52	.71			.53
What (N)	.64			.45	.65			.47
What (V)	.61			.44	.62			.45
When	.64			.46	.65			.47
Where	.67			.48	.68			.50
Which	.71			.55	.72			.57
Who	.67			.49	.68			.51
Why	.67			.51	.68			.52
Test Wiseness:			.36	.20			.36	.20
Short Form Test of Academic Aptitude:								
Language IQ			.85	.81			.86	.81
Non-Language IQ			.84	.76			.84	.76
California Achievement Test:								
Words in Context	.47	.47	.54	.73	.45	.47	.56	.74
Language Mechanics	.35	.68	.31	.69	.34	.68	.33	.69
Standard English		.60		.49		.61		.49
Sentence Structure		.67		.49		.68		.50
Sentence Parts and Functions		.68		.48		.68		.48
Transformations		.66		.44		.66		.44
Comprehension Item Classification								
Lit.Comp./Pass. Depend.	.44	.38	.49	.57	.43	.38	.50	.58
Lit. Comp./Pass. Indep.	.45	.40	.47	.58	.44	.40	.48	.59
Non-Lit. Comprehension		.43	.49	.48		.43	.50	.48
Eigenvalue	10.15	1.77	1.17		9.18	1.76	1.17	
Cumulative Proportion of Total Variance	.42	.50	.55		.42	.50	.55	

38