

Resource Consultant Training Program
Research Report No. 9

RCTP

Pilot-Test Results on
Written Retell:
An Ecological
Curriculum-Based
Measure for Secondary
Students

Richard Parker
Gerald Tindal

Published by
Resource Consultant Training Program
Division of Teacher Education
College of Education
University of Oregon

Copyright © 1989 University of Oregon. All rights reserved. This publication, or parts thereof, must not be used or reproduced in any manner without written permission. For information address University of Oregon, Department of Education, Resource Consultant Training Program, 275 Education, Eugene, OR 97403-1215.

Parker, Richard
Tindal, Gerald
Pilot-Test Results on Written Retell: An Ecological Curriculum-Based Measure for Secondary Students
Research Report No. 9

Staff

Gerald Tindal, Program Director
Jerry Marr, Editor
Clarice Skeen
Denise Styer
Mark Baldwin

Acknowledgements

Preparation of this document was supported in part by the U.S. Department of Education, grant number G008715106. Opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Cover design: George Beltran

PILOT-TEST RESULTS ON WRITTEN RETELL: AN ECOLOGICAL CURRICULUM-BASED MEASURE FOR SECONDARY STUDENTS

Richard Parker
Gerald Tindal

Abstract

Curriculum-based measurement (CBM) has been advocated by a number of professionals and practiced in many districts, but writing on the subject and its application has focused on elementary schools. This study addresses the adoption of CBM in secondary settings. A written retell measurement task was used to sample performance in a functional task necessary for success in content-area classes. Students from three grade levels (6, 8, and 11) were assessed using written retell along with two other criterion measures: a Maze comprehension task and a creative writing sample. The retell measure was scored with several metrics, and the data analyzed to ascertain changes across grade levels and relationships among different metrics. The findings are interpreted as supportive of using curriculum-based measurement formats in secondary settings.

INTRODUCTION

As students progress through upper intermediate and middle grades to secondary school, an increasing proportion of their assignments requires the integrated application of several basic language arts skills. While spelling, handwriting, written composition, reading fluency, and reading comprehension are often separate lesson foci in the early grades, they more often are applied together in later grades. At the same time, secondary students begin to approach academic tasks more globally, integrating once-separate skills into the functional goal of assignment completion (Birnbaum, 1982; Langer, 1986).

Concurrent with the increased integration of basic skills in lessons and assignments, a second major change in instructional focus occurs as students move through the grades. The content areas of science, social studies, literature, health, home economics, and vocational education assume greater importance. Less class time is spent on

instruction in language arts skills, and more is dedicated to increasing content area knowledge and accessing content area information (Halpern & Benz, 1987). Basic skills such as reading become tools for exploring subject matter and completing functional tasks (Chall, 1983). Reading becomes important as a learning medium rather than a subject area (Berryhill, 1984; Cheek & Cheek, 1983; Herber, 1978; Niles, 1985; Vacca, 1981).

Unfortunately, most secondary special education programs do not reflect these foci of secondary school curriculum and instruction; services for students with learning problems are segregated from regular class curricula (Halpern & Benz, 1987) and dedicated to basic skills instruction (Houck, Geller, & Engelhard, 1988). The value of these programs is, however, being strongly challenged (Readence, Bean & Baldwin, 1985; Reynolds, Wang, & Walberg, 1987). A case has been made that instruction in basic skills for students with mild

handicaps is less cost-effective in the higher grades, and that the focus should shift to study skills or strategies that may be immediately applied to the successful completion of classroom assignments (Deshler, Schumaker, & Lenz, 1984).

Several innovative programs have emerged in recent years where instruction is centered on support in the regular classroom, through study skills/strategies instruction, tutorial assistance in regular class assignments, and contractual agreements with the regular classroom teacher to modify course expectations (Adams, Carnine, & Gersten, 1982; Lovitt, Rudsit, Jenkins, Pious, Benedetti, 1986; Schumaker & Deshler, 1984; Tindal, Shinn, Walz & Germann, 1987). This type of instruction is proposed for *all* students by many leaders in the field of reading. For example, Anderson, Hiebert, Scott, & Wilkinson (1985) state that the "most logical place for instruction in most reading and thinking strategies is in social studies and science rather than in separate lessons about reading" (p. 73).

The re-conceptualization of secondary special services delivery for students with mild handicaps has been fortuitously paralleled by a re-conceptualization of how basic skills should be assessed, mainly due to the influence of cognitive researchers. Valencia and Pearson (1988) argue for reading comprehension assessment that is integrated and applied, rather than subskill-specific:

If we assess the skills rather than the orchestration of them, we run the risk of monitoring and teaching reading skills at an inappropriate level of specificity; that is, we might come to hold students and ourselves accountable for isolated, discrete behaviors that might never really result in comprehension of text. ... Reading assessments must have ecological validity; they must deal with reading tasks that reflect school and lifelike situations. (p. 29)

Integrated reading/writing assessment thus would be consistent with an increasing body of literature that reflects common cognitive processes, strategies, and skills in these two performance areas (Aulls, 1985; Birnbaum, 1982; Tierney & Leys, 1986; Tierney & Pearson, 1983).

In the elementary grades, curriculum-based measurement (CBM) has proved useful for (a) identifying students who require special services (Shinn, Tindal, & Stein, 1988), (b) ongoing progress monitoring for making formative instructional decisions (Deno & Fuchs, 1987; Fuchs, 1986), and

(c) mainstreaming students who had been placed in separate special education programs (Germann, 1985). Part of the power of curriculum-based measures derives from their content validity, that is, their close match with the instructional objectives, presentation and response modes, and materials used in the elementary grades (Deno, 1985; Idol, 1986; Tucker, 1986)—a match that is unlikely with standardized reading tests (Good & Salvia, 1988; Jenkins & Pany, 1978; Shapiro & Durr, 1987).

In secondary grades, with the increased requirement for integrated application of multiple language arts skills and for understanding and recalling content information, the traditional elementary CBM measures may be less content valid. Whether the purpose of assessment is identification for special services, mainstream programming, or progress monitoring, assessment tools are required in secondary special education with greater content validity than traditional CBMs, which treat basic skills in isolation, for example, "correct word sequences" in writing (Videen, Deno, & Marston, 1982), and "rate of words read correctly" in reading (Deno, Mirkin, & Chiang, 1982). Measures are needed that (a) better fit the regular classroom and the integrated, supportive service options existing in model programs, (b) include both integration of basic skills and accessing and recalling of content area information in a classroom-relevant task, and (c) assess products with functional importance in secondary classrooms—"assignments."

Written retell procedures accomplish all three functions, since they involve both reading comprehension and written expression applied to content area material, and they are a primary activity in secondary classrooms. Students silently read a designated passage from a classroom text and immediately display their recall of the content through a written summary. While *oral* retell as an assessment measure has a rich research history of over 15 years (Calfee & Drum, 1986; Carroll & Freedle, 1972), the research base on *written* retell is much more sparse, although supportive.

Theoretical researchers of text structure (Kintsch & vanDijk, 1978; Meyer, 1977) have found students' written summaries of reading assignments useful in revealing their levels of comprehension. Several researchers have demonstrated that practices such as writing content summaries of text passages positively affect both the skills and content learning (Doctorow, Wittrock, & Marks, 1978; Howard, 1983; McGee & Richgels, 1985; Taylor & Beach, 1984). The method, format, and

time allowance for writing has not been standardized across studies, however. In some cases, summaries are written after every paragraph; in other cases, outline notes are taken, and then summaries are written from the notes (see Niles, 1985 for a review). Students may or may not be allowed to look back at the story while writing.

Even with these procedural variations, however, written retell appears to be supported by research as both an instructional and assessment tool. Jenkins, Heliotis, Stein, and Haynes (1987) demonstrated that a post-reading written retell procedure completed by 32 students with learning disabilities significantly improved content recall. Furthermore, the strategy generalized to other, similar tasks. Fuchs, Fuchs, and Maxwell (1988) used written retell, along with a number of other measures (Stanford Achievement Test, Cloze, short-answer multiple-choice questions, correct reading rate) to assess the reading comprehension of 70 mildly to moderately handicapped junior high school students. *Written* retell scores correlated significantly higher with the Stanford Reading Comprehension subtest, the multiple choice questions, the Cloze test, and correct reading rate, than did *oral* retell scores. Fuchs, Fuchs, & Maxwell (1988) conclude that of those curriculum-based reading comprehension measures investigated, written retell may be the most feasible for classroom use.

Although the written retell task is considered to have high content validity for secondary classrooms, its measurement properties are as yet unknown and must be investigated prior to use by teachers for important instructional decisions. This study seeks to provide two types of information: (a) normative performance on written retell indices, and (b) the relationship between written retell indices and external criterion measures. The first type of information is needed to anticipate how written retell might be used as a decision-making tool within school settings. Can it be used to distinguish students at different grade levels? What range of performance can be expected within regular classrooms? Those questions which relate more to the adequacy of a test for progress monitoring (e.g., sensitivity to change) are not addressed in this study. Here, we are only considering the utility of written retell as a screening instrument for classification decisions, requiring less precision.

The second, criterion-related type of information provided by this study will help explicate written retell as a construct by embedding retell

indices in a "pattern of empirical relationships between test scores and other variables" that "clearly indicates the meaning of the test score." (American Psychological Association, 1985, pp. 9-10). The resulting data will help answer practical questions for the practitioner, such as whether written retell indices can be used to replace or supplement common classroom measures, such as oral reading fluency or a Cloze test. Specifically, the study will provide normative and criterion information by addressing four questions:

1. What is the distribution of performance by students in regular classrooms on written retell tasks involving grade-level reading materials?
2. Do students generally improve in performance on written retell tasks over a six year period (Grade 6 to Grade 11)?
3. How do written retell indices relate to separate reading comprehension (Maze) and written expression tasks which do not require recall of content?
4. Do the relationships among reading/writing/written retell skills change or remain stable over the 7-year period?

METHOD

Population

Two hundred sixty-seven students were selected at three grade levels (Grade 6: 95, Grade 8: 97, Grade 11: 75) from a school district located in a lower-middle SES community located on the west coast. Students were obtained by randomly selecting four classrooms at each grade level, and testing all those present on the day of test administration.

Measures

Grade 6 and 8 students completed three tasks—a Maze test, creative writing from a story starter, and written retell—while Grade 11 students completed all but the Maze. The Maze and written retell were based upon a single reading passage at each grade level, with readability matching the instructional level for that grade. The reading passage, along with a story starter, also set the topic and served as background material for the creative writing task. The Maze and creative writing tasks served as independent, curriculum-based measures of reading and writing, respectively, against which the retell scores were compared.

Reading

All passages were approximately 250 words in length. For sixth graders, the passage was sampled from a social studies text and dealt with a description of Greece. The passage for eighth graders was

sampled from a biological science text and dealt with a description of the skeletal/muscular system. For 11th graders, the passage was drawn from a voter information pamphlet for a proposed bill mandating helmet use by motorcyclists. (See Appendix A for texts of passages.)

Maze

The first measure to be administered, a "multiple choice Cloze" or Maze test (Hinofotis & Snow, 1977; Howell & Kaplan, 1981), was produced from the Grade 6 and Grade 8 passages. Unlike the Cloze, the Maze test does not require a word-writing response from students, and scores of 85-90% correct are expected of an able reader, in contrast to only 45% for the Cloze. In producing the Maze test from the 250-word passages, the first and last sentences were left intact and every sixth word omitted for the remaining text. The resulting approximately 35 omitted words formed the pool or universe of distractors. For each blank in the Cloze test, students were presented with a five-option, multiple choice "keyword search" task (Roid & Haladyna, 1982). The four distractors for each test item were selected at random from the pool of distractors. The only stipulation to the random selection procedure was that if a randomly selected distractor happened to be a "reasonable answer," it was replaced with another selection.

The resulting Maze test was produced on a single page, with the story on the left of the page and the five response options on the right (see Appendix B). Each blank in the story was keyed with a numeral to match the five options on the right. Students were asked to circle the best answer among the five options.

Creative writing

After reading the selected passage, and completing the Maze and retell, students were presented with a creative writing task, following the procedures outlined by Marston, Lowry, Mirkin and Deno (1981). A story starter was presented first—a short, written scenario based on the content of the preceding reading selection. Students then were directed to be creative in writing a response and encouraged to make use of information from their respective reading passages within their compositions. Sixth graders were asked to describe a Grecian travelogue (after reading about "Ancient and Modern Greece"), eighth graders described the creation of a human-like robot (after reading "The Skeletal and Muscular Systems"), and 11th graders critiqued legislation for eliminating "cruising" by teenagers (after reading "Motorcycle Helmet Law").

The creative writing samples were rated holistically according to their "communicative effectiveness," scaled from 1 (very poor) to 5 (very effective), with no intermediate descriptors. Raters were trained undergraduate students in education who were assisted in the scoring task by the following definition of good writing, produced by consensual agreement among a team of four practicing special education teachers:

Good writing clearly communicates to the reader the ideas/story of the writer. Good writing requires legible handwriting or printing, as well as distinguishable words, phrases, and sentences. Coherent linking of ideas from one sentence to the next also contributes to good writing. (Hasbrouck, 1987, p. 2)

This procedure had been used extensively by the research team in previous studies, and typically produces indices of inter-judge agreement from .75 to .90 for student writing at various grade and ability levels (Tindal & Parker, 1989). *Within-grade* interrater reliability for holistic judgments in this study was $r = .76$. Higher reliabilities (around $r = .83$) have been obtained by the authors when holistic judgments are made on writing samples pooled across grade levels. Selection of the appropriate reliability will depend on whether decisions from test data are made within or across classroom/grade level groups.

Retell

Following completion of the Maze test, students were instructed to read the passage and write down in an organized manner—using sentences and paragraphs—as many main ideas and details as they could recall from the text, without looking back at the story. They then read the assigned passage, noted elapsed time, turned the page over, completed a written retell task on lined paper printed on the back of the passage, and noted total elapsed time.

Retells were scored using three measures: (a) number of words written (Tot.Wds), (b) holistic judgment of communicative effectiveness (Retell.Hol), and (c) number of passage-related idea units written (Recall). The word production count employed the procedures developed by Deno, Marston, and Mirkin (1982). To judge communicative effectiveness holistically, the same procedures were used for both the retell and creative writing samples. For recall scoring, the sentence clause (dependent or independent) was se-

lected as the smallest unit of content analysis, a method somewhat coarser than the more commonly used "pausal unit" scoring (Clark, 1982). However, it proved easier and faster to use than pausal units, a major concern in developing scoring procedures for routine classroom use. A student's writing sample was matched against the idea contained in each original passage clause, and the student was given 1 or 2 points, respectively, for partial or full accurate restatement of the sentence idea. Mis-statements were ignored for this study. When content retold by a student matched more than one original sentence clause (i.e., the passage repeated an idea), the student was given credit under only one clause—the one occurring first in the passage. For this study, there was no attempt to weight the individual content clauses by importance. (See Appendix C for examples of retell scoring sheets.)

Interscorer reliability of the number of passage-related idea units written (Recall) was $r = .78$ within a single grade-level cohort of 73 cases. Across a sample of 68 cases, composed of equal proportions of grade six and eight samples, the interrater agreement was greater, $r = .81$. Brown (1983) provides a good discussion of the size of reliability estimates due to within-grade *versus* cross-grade sampling. Nunnally (1978) indicates that reliability coefficients of .70 to .80 often suffice in "early stages of research on predictor tests," where the principal concern is with mean differences between groups (p. 245). Where measures are used to make important placement decisions for individual students, however, a reliability of .90 is minimal, and .95 is "the desirable standard" (Nunnally, 1978, p. 246).

Procedures

All measures were group-administered during one class period by two trained supervisors who had extensive experience in curriculum-based assessment. The following order of administration was used to avoid confounding outcomes from the different measures: (a) Maze, (b) silent reading, (c) written retell, and (d) creative writing. Students were told the results would not be shared with their teachers and their performance on the tests would not have any bearing on course grades. (Teacher Directions appear in Appendix D.)

The Maze test was completed in less than 15 minutes. Next, students completed the silent reading and retell in a single session. Students were directed to (a) read the passage, (b) record the time elapsed from the start (which was displayed in 15-

second intervals on the blackboard), (c) turn the page over to write a summary of the content as thoroughly as they could without looking back at the passage, and when completed, (d) write down the elapsed time. A 15-minute limit was established, with a 13-minute warning delivered as a prompt to finish writing.

The creative writing task was administered last. Students were read the "story-starter" paragraph and then directed to create a story that best completed it. They were warned that only 10 minutes would be available and that they were to make a slash on the line where they were writing at the 3-minute mark.

Summary Scores and Analyses

From the three tasks (Maze, written retell, and creative writing), seven summary scores were obtained. The Maze yielded a "number correct" score of the approximately 35 multiple-choice questions (Maze). The creative writing samples yielded holistic ratings of writing quality (communicative effectiveness) on a 1 to 5 scale (Crt.Hol). From the written retell, five scores were derived: silent reading rate (Rdg.Rate), written retell rate (Rtl.Rate), total number of words written (Tot.Wds), holistic ratings of the writing quality of the written retell (Retell.Hol), and amount of reading content accurately recalled (Recall).

Summaries of central tendency, dispersion, and distribution shape were tabulated for all seven indices, and notched box plots (Cleveland, 1985; Chambers, Cleveland, Kleiner, & Tukey, 1983) were produced for all but the holistic ratings. Tests of differences between Grade 6, 8, and 11 level means were also conducted. The seven indices were then inter-correlated at each grade level. Finally, hierarchical cluster analysis was applied to the correlation matrices to help interpret skill clusters and changes in these clusters over the six years of instruction, Grades 6 to 11.

RESULTS

Descriptive Data

One of the most efficient summary displays of a distribution is Tukey's (1977) box plot, which shows the inter-quartile range (25th to 75th percentiles) with the median marked by a horizontal line. Box plots allow the skewness of a distribution to be discerned, since extreme scores are not hidden but plotted as individual small circles. Finally, the notches around the median, representing 95% confidence intervals, allow an estimate of whether the medians of two distributions are statistically

different in size. If two adjacent distributions have overlapping notches, we *cannot* say with 95% confidence that the true medians are different (Chambers, Cleveland, Kleiner, & Tukey, 1983).

Box plots were produced (Figure 1) for all indices but the holistic ratings, which would be uninformative since they were conducted *within*, not *across*, grade levels, using a "floating standard" based upon the group of writing samples at each grade level. In a few cases, extremely high or low individual scores were cut out of the box plots to reduce the plots' overall dimensions and magnify the size of individual boxes.

A cursory glance at the box plots shows relatively little difference between most Grade 6 and 8 median scores. On the other hand, all five indices except the Maze show significant differences between Grade 11 median scores and those for the lower two grades. These results were largely confirmed through more traditional parametric analy-

ses. Anovas performed on the five sets of scores from the three grade levels revealed no significant differences between Grades 6 and 8; Grade 11 scores were significantly different (Tukey, $p < .05$) from each of the other two grades, however, for four of the scores: Rtl.Rate, $F(2,263) = 29.72$, $p < .0001$; Tot.Wds, $F(2,263) = 32.63$, $p < .0001$; Recall, $F(2,263) = 7.14$, $p < .001$; Retell.Hol, $F(2,263) = 5.41$, $p < .005$.

The box plots also indicate (from the vertical axes) a wide range in raw scores within any one grade level on most measures. The size of most scores within any grade level varied by a factor of 10. The recall scores also demonstrate this large range. The increasing heights of boxes in the plots for Reading Rate and Retell Rate also indicate that, within classroom groupings, the spread of skills becomes greater as students get older. This effect is more pronounced for reading than for writing rate.

Figure 1. Interquartile Box Plots for Five Measures

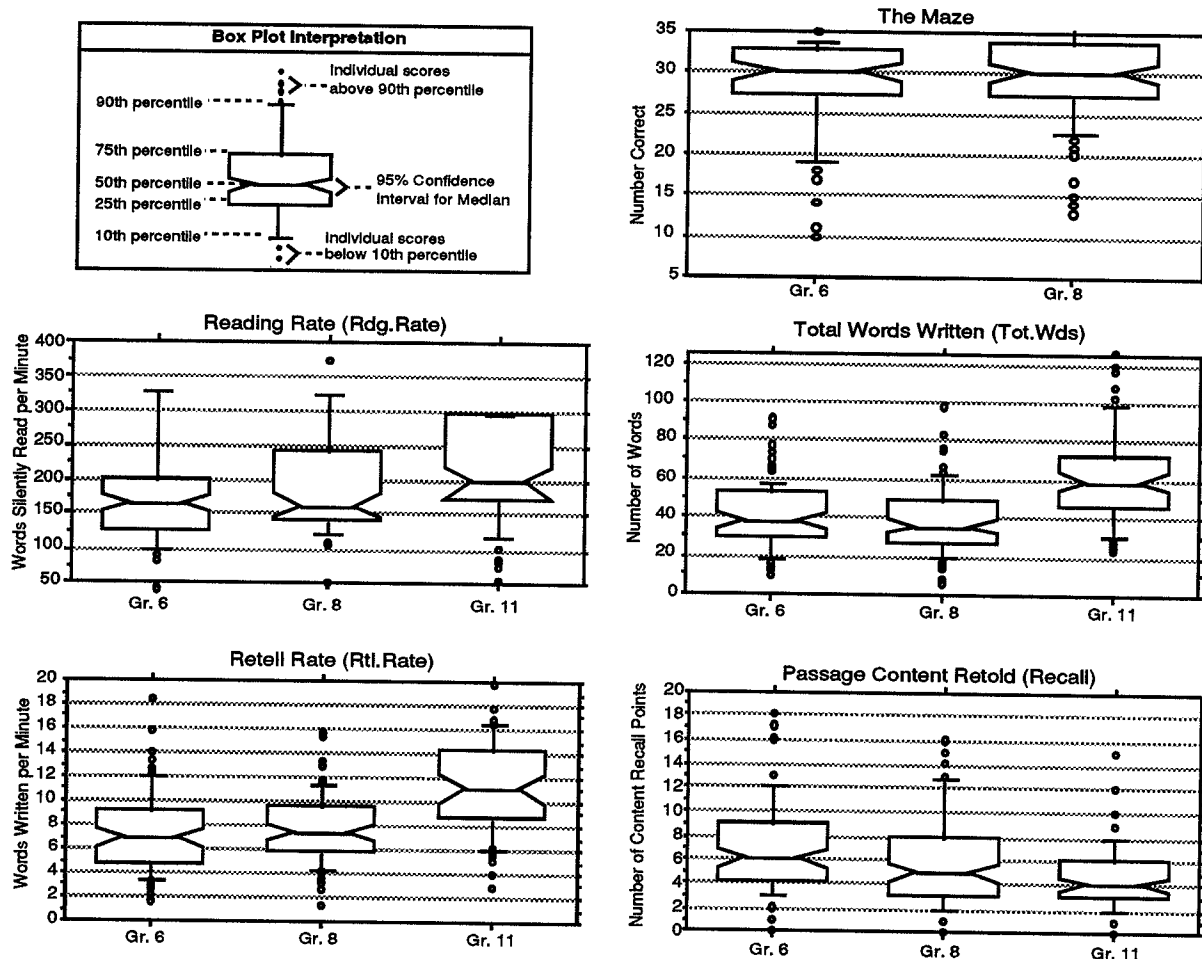


Table 1. Descriptive Statistics for Grades 6, 8, and 11 on Three Curriculum-Based Literacy Tasks: Maze Test, Creative Writing, and Written Retell

Written Retell							
<u>Grade 6: (n = 95)</u>							
	<u>Maze</u>	<u>Crt.Hol</u>	<u>Rdg.Rate</u>	<u>Rtl.Rate</u>	<u>Tot.Wds</u>	<u>Recall</u>	<u>Retell.Hol</u>
Min	10.0	1.0	41.0	1.7	10.0	0.0	1.0
Max	35.0	5.0	490.0	18.4	91.0	18.0	5.0
Mean	28.5	2.5	179.6	7.3	39.7	6.6	2.7
SD	5.6	0.7	89.4	3.3	17.2	3.7	0.8
Skewness	-1.3	-.29	1.61	.66	.54	.81	.18
<u>Grade 8: (n = 97)</u>							
	<u>Maze</u>	<u>Crt.Hol</u>	<u>Rdg.Rate</u>	<u>Rtl.Rate</u>	<u>Tot.Wds</u>	<u>Recall</u>	<u>Retell.Hol</u>
Min	13.0	1.0	50.1	1.2	6.0	0.0	1.0
Max	35.0	5.0	500.7	15.7	98.0	25.0	5.0
Mean	29.3	2.7	204.3	7.7	39.1	5.9	2.7
SD	5.1	0.7	96.9	2.9	17.8	4.2	0.8
Skewness	-1.08	-.26	1.78	.38	.79	1.45	-.07
<u>Grade 11: (n = 75)</u>							
		<u>Crt.Hol</u>	<u>Rdg.Rate</u>	<u>Rtl.Rate</u>	<u>Tot.Wds</u>	<u>Recall</u>	<u>Retell.Hol</u>
Min	-	1.0	53.0	2.9	24.0	0.0	1.0
Max	-	5.0	588.0	27.4	127.0	15.0	5.0
Mean	-	2.8	220.6	11.4	60.9	4.4	3.1
SD	-	0.8	113.2	4.3	23.1	2.7	0.7
Skewness	-	-.06	1.85	.89	.76	1.12	-.09

Maze = Maze Test

Crt.Hol = Creative Writing Task: Holistic Judgment of 'Communicative Effectiveness'

Rdg.Rate = Written Retell Task: Reading Rate

Rtl.Rate = Written Retell Task: Writing Rate

Tot. Wds = Written Retell Task: Total Number of Words Written

Recall = Written Retell Task: Number of Content Recall Points

Retell.Hol = Written Retell Task: Holistic Judgment of 'Communicative Effectiveness'

For retell scores, the opposite effect is noticed; the range and spread of scores decreases as students reach Grade 11. Retell also shows a surprising overall decrease in scores from one grade to the next. The Maze is noteworthy for the consistency in scores by students in the grade-level material; roughly 80% correct is the average score at both grades. These score ranges and dispersions are also reflected in the descriptive statistics, presented in Table 1, above.

Skewness of scores reflects deficiencies in tests, such as ceiling or floor effects, that indicate a need to adjust task item difficulty. Highly skewed scores are unequally sensitive to performance differences along the achievement continuum and do not lend themselves well to parametric statistical analysis (Hays, 1981). The box plots indicate skewed distributions through inequality of the upper and lower parts of the box, and unequal lines or "stems" attached to the top and bottom of the box. The Reading Rate (Rdg.Rate) metric shows the most

highly skewed distributions—all in a positive direction. Furthermore, this skewness increases from one year to the next (see Table 1). In other words, score distributions increasingly “bunch up” at the lower end (100-175 WPM), and spread out at the upper end (200-300 WPM). The Passage Content Retell scores also show positive skewness, with scores “bunched up” toward the floor (2 to 4 points earned out of a possible 50). The Maze distributions also show pronounced skewness, but in a negative direction; a ceiling effect is noticeable, as students “bunch up” against the 100% limit (35/35 correct).

To accurately compare the skewness scores from Table 1 across measures and grade levels, they must first be standardized by the cubes of their respective standard deviations (Ghiselli, Campbell, & Zedeck, 1981). After standardizing by SDs, Recall scores proved most highly skewed, followed by Retell Rate scores.

Although the standard errors of mean (SEM) are reflected in the notched box plots, standard errors of measurement (SEM) are required to estimate the confidence intervals around individual student scores. For input into the formula, $SEM = s_x \div 1 - R_{xx}$, reliability estimates (e. g., internal consistency, parallel forms, test-retest) are selected on the basis of “the sources of error that are of the greatest concern” (Brown, 1983, p. 96). Because error is additive, the most conservative estimates include several potential sources of error, for example stability over time of alternate forms, scored by different raters (Brown, 1983).

Here, the computation of SEM for content recall (Recall) is based on inter-scoring agreement only, as stability and retest data were not available; it should therefore be considered only as an *upper limit* to reliability. Using the within-grade inter-scoring reliability of $r = .78$ yielded $SEM = 2$ for Grade 8 recall scores. At the 95% level of confidence, scores at the 75th and 90th percentiles can be differentiated (barely), but scores at the 10th and 50th percentiles cannot.

Correlations and Cluster Analyses

At the Grade 6 level, cluster analysis of the correlation matrix (as shown in Figure 2) yielded four clearly identifiable clusters: (a) the Maze Test (Maze), (b) holistic rating of creative writing (Crt.Hol), (c) reading rate (Rdg.Rate), and (d) the four written recall variables—retell rate (Rtl.Rate), number of words written (Tot.Wds), content recall (Recall), and holistic rating of the retell writing sample (Retell.Hol). The most closely related

group of variables ($r = .68$ to $.78$) are Tot.Wds, Recall, and Retell.Hol—all three based upon the written retell task. The variable Rtl.Rate is an outlying member of this group, related moderately to each of the other three ($r = .49$ to $.66$). The variable Rdg.Rate is isolated, substantially related only to Rtl.Rate ($r = .45$). Crt.Hol and the Maze test are very isolated, unrelated to each other ($r = .17$) or to other variables.

Grade 8 scores are similar to Grade 6, with four clusters also apparent (see Figure 3). The tight cluster ($r = .65$ - $.79$) of Tot.Wds, Recall, and Retell.Hol does still exist. Rtl.Rate has been replaced by the Maze test, however, as a fringe member ($r = .10$ - $.59$) of the central cluster, most closely related to Tot.Wds ($r = .59$). As at the Grade 6 level, in Grade 8 Rdg.Rate and Crt.Hol are among the most isolated variables ($r = -.21$ - $.28$).

A cluster analysis solution makes use of all information in the correlation matrix, as does factor analysis, and is influenced by adding or omitting particular variables. The Grade 11 cluster tree is therefore not strictly comparable with those for the other two grades because Grade 11 students were not administered the Maze test (see Figure 4). For them, the cluster of Tot.Wds, Recall, and Retell.Hol still exists, but is not as consistent ($r = .33$ - $.58$) as it was for the two earlier grades. Rtl.Rate continues to be the fringe member ($r = .23$ - $.57$) of this group, most strongly related to Tot.Wds ($r = .57$).

DISCUSSION

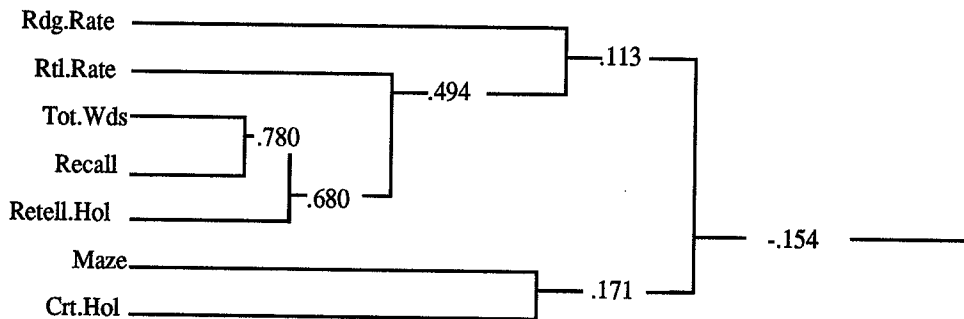
The purpose of this study was to explore certain objectively scorable facets of the written retell task, applied to content-area reading material, at Grades 6, 8, and 11. This task was selected for two reasons. First, it appears to be content- and context-valid (i.e., similar to routinely-occurring tasks in secondary level classrooms). Second, it integrates and applies the two basic skills of reading comprehension and written expression to content area learning. It has been argued that student assessment through this complex, classroom-valid task may provide more and different information than that obtained through separate assessments of component skills.

More specifically, this study examined various written retell indices to provide normative and criterion-related answers to four main questions: (a) What is the distribution of performance by students in regular classrooms on written retell tasks involving grade-level reading materials; (b) do students generally improve in performance on written retell tasks over a 6-year period (Grade 6 to

Figure 2. Correlation Matrix and Related Cluster Analysis for Seven Scores from Three Curriculum-Based Literacy Tasks: Maze Test, Creative Writing, and Written Retell: Grade 6 ($n = 95$)

	<u>Maze</u>	<u>Crt.Hol</u>	<u>Rdg.Rate</u>	<u>Rtl.Rate</u>	<u>Tot.Wds</u>	<u>Recall</u>
Crt.Hol	.171
Rdg.Rate	.053	-.154
Rtl.Rate	.150	-.010	.455	.	.	.
Tot.Wds	.176	.034	.252	.669	.	.
Recall	.196	.026	.135	.494	.780	.

Hierarchical Cluster Analysis*, with Standardized Joining Distances



- Maze = Maze Test
- Crt.Hol = Creative Writing Task: Holistic Judgment of 'Communicative Effectiveness'
- Rdg.Rate = Written Retell Task: Reading Rate
- Rtl.Rate = Written Retell Task: Writing Rate
- Tot. Wds = Written Retell Task: Total Number of Words Written
- Recall = Written Retell Task: Number of Content Recall Points
- Retell.Hol = Written Retell Task: Holistic Judgment of 'Communicative Effectiveness'

*Using Complete Linkage Method

Grade 11); (c) how do written retell indices relate to separate reading comprehension (Maze) and written expression tasks which do not require recall of content; and (d) do the relationships among reading/writing/written retell skills change or remain stable over the 6-year period? These questions were addressed using five objective scoring indices: reading rate (Rdg.Rate), written retell rate (Rtl.Rate), total number of words written (Tot.Wds), a holistic judgment of the writing quality of the written retell (Retell.Hol), and amount of reading content accurately recalled (Recall).

The first question focused on the performance distributions of the five written retell indices. Content recall proved to be a difficult task; most students were able to recall and relate in writing only roughly 10% of the roughly 25 details and main ideas identified in each reading passage. The

majority of scores at each level “bunched up” at the low level of 3 to 4 points, indicating only one or two completely and accurately paraphrased clauses from the passage.

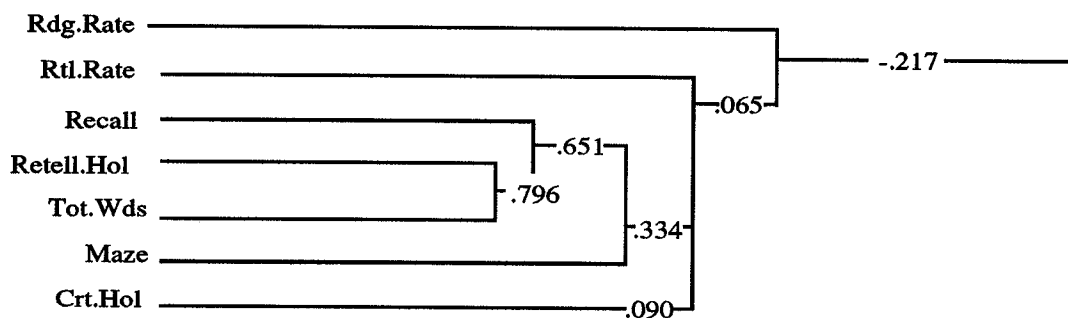
The highly skewed recall scores ranged widely (by a factor of 15 to 25) within each grade level, compared to tenfold within-grade score ranges for silent reading rate, writing rate, and total number of words written. In the allotted written retell time, most students wrote about 40 words (five sentences) in Grades 6 and 8 and 60 words (6 to 7 sentences) in Grade 11. It is clear that the low recall scores were not due primarily to lack of writing time, or even lack of amount written. Most words and sentences were clearly irrelevant to the passage.

Although the recall difficulty at the Grade 11 level may have been a function of difficulty in interpreting the voter information pamphlet, this

Figure 3. Correlation matrix and related cluster analysis for seven scores from three curriculum-based literacy tasks: Maze test, creative writing, and written retell: Grade 8 ($n = 97$)

	<u>Maze</u>	<u>Crt.Hol</u>	<u>Rdg.Rate</u>	<u>Rtl.Rate</u>	<u>Tot.Wds</u>	<u>Recall</u>
Crt.Hol	.180
Rdg.Rate	-.050	-.217
Rtl.Rate	.065	.088	.053	.	.	.
Tot.Wds	.334	.284	-.129	.597	.	.
Recall	.408	.090	-.117	.105	.651	.
Retell.Hol	.395	.239	-.123	.370	.796	.718

Hierarchical Cluster Analysis*, with Standardized Joining Distances



Maze = Maze Test

Crt.Hol = Creative Writing Task: Holistic Judgment of 'Communicative Effectiveness'

Rdg.Rate = Written Retell Task: Reading Rate

Rtl.Rate = Written Retell Task: Writing Rate

Tot. Wds = Written Retell Task: Total Number of Words Written

Recall = Written Retell Task: Number of Content Recall Points

Retell.Hol = Written Retell Task: Holistic Judgment of 'Communicative Effectiveness'

*Using Complete Linkage Method

explanation cannot account for performance at Grades 6 and 8. For them, the Maze scores reflect generally strong passage understanding (85% correct). Rather, it is likely that recalling information from text and relating it in writing are skills distinctly different from those directly measured by the "enabling" skills of reading rate, writing rate, writing quantity, passage comprehension, and general writing quality. In short, the whole appears to be more than the sum of its parts.

The reliability of individual recall scores, as reflected in confidence intervals placed around Grade 8 scores, was insufficient to make even gross distinctions of 40 percentile points with reasonable certainty below the median. Thus, it is premature to use the recall index as a screener for special or remedial services.

The second question posed by this study was whether students generally improve in performance on written retell (with grade-level reading material) from Grade 6 to Grade 11 (as oral reading rate) or remain stable (as Cloze or Maze tests). Bearing in mind that the difficulty of the written retell tasks could only loosely be controlled by readability formulae and informal judgments, the Maze tests constructed for Grades 6 and 8 produced nearly identical score distributions. Retell scores, on the other hand, dropped steadily from one grade to the next—contrary to any other score index. This disturbing finding reinforces the hypothesis above that written recall is a combination of skills uniquely different from the component skills measured. From this finding we also infer that this unique aggregate of skills does not steadily improve over the later school years.

It was also predicted that the holistic judgments of "communicative effectiveness" on a related creative writing task would correlate highly with the written retell objective indices. This finding failed to appear at any of the three grade levels. Even the same judgments of communicative effectiveness made on the two writing tasks were unrelated. These results caution us against over-generalizing about a student's "writing skill" across types of writing tasks!

In summary, there appears to be a composite of three or four moderately interrelated skills that (a) can be objectively measured from a written retell task and (b) appears to be uniquely different from reading comprehension skills, as measured by a Maze test, or writing quality, as measured by a creative writing task from story starter. Furthermore, reading and writing rates (in words per minute) are peripheral to the composite of written retell skills, which appear to become less interrelated as grade level increases.

These generalizations strongly support the need to pursue increasingly complex, functional, classroom-valid language arts measures at the middle and secondary school levels. Curriculum-based measures that have proved valuable in elementary grades (Tindal, Germann, & Deno, 1983) may not be as useful to document secondary classroom performance. Written retell-based indices such as "number of words written," "content accurately recalled," and "communicative effectiveness of the writing sample" from written retells, may better serve as measures of requisite skills for regular class success at the secondary level.

The reliabilities reported for "content accurately recalled" in writing are not sufficiently high to allow individual placement decisions, and are certainly insufficient for the use of the written retell task in progress monitoring. At this point in its psychometric career, written retell is recommended mainly for informal classroom use as a test with superior content-validity, which measures skills not tapped by other, more simple skills tests. Neither do reliability estimates of the content recall score warrant its present use as a screening tool for special or remedial services. At present, increased content validity appears to have been purchased at an excessive cost to reliability.

Future research should focus on systematic manipulation of elements of the recall task to reduce the number of potential sources of unwanted score variation. Scoring formats or guidelines and training will certainly need to be improved to increase inter-scoring reliability. Standardized read-

ing passages may need to be selected which especially lend themselves to retell. Finally, more structuring of the student's writing task may be required to reduce response variation in extraneous dimensions. Hopefully, these steps will yield a task still with high content and context validity, and with reliability sufficiently high for classroom and school decision-making.

REFERENCES

- Adams, A., Carnine, D., & Gersten, R. (1982). Instructional strategies for studying content area texts in the intermediate grades. *Reading Research Quarterly*, 18(1), 27-55.
- American Psychological Association, AERA, NCME (1985). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the commission on reading*. Washington, DC: Department of Education.
- Aulls, M. (1985). Understanding the relationship between reading and writing. *Educational Horizons*, 64, 39-44.
- Berryhill, P. (1984). Reading in the content area of social studies. In M. M. Dupuis (Ed.), *Reading in the content areas: Research for teachers* (pp. 66-81). Newark, NJ: International Reading Association.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). New York: Holt, Rinehart & Winston.
- Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher*, 10(2), 14-24.
- Birnbaum, J. (1982). A study of reading and writing behaviors of selected fourth and seventh grade students. *Research in the Teaching of English*, 16, 241-261.
- Calfee, R., & Drum, P. (1986). Research on teaching reading. In M. C. Wittrock (Ed), *Handbook of research on teaching: Third edition*. (pp. 804-849). New York: Macmillan.
- Carroll, J. B. & Freedle, R. O. (Eds.) (1972). *Toward a literate society*. New York: McGraw-Hill.
- Chall, J. (1983). *Stages of reading development*. New York: McGraw Hill.
- Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphic methods for data analysis*. Boston: Duxbury Press.
- Cheek, E. H., & Cheek, M. C. (1983). *Reading instruction through content teaching*. Columbus, OH: Charles E. Merrill.

- Clark, C. H. (1982). Assessing free recall. *The Reading Teacher*, 35(4), 434-439.
- Cleveland, W. S. (1985). *The elements of graphing data*. Monterey, CA: Wadsworth Advanced Books and Software.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem-solving. *Focus on Exceptional Children*, 19, 1-16.
- Deno, S. L., Marston, D., & Mirkin, P. K. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children*, 48, 368-371.
- Deno, S. L., Mirkin, P., & Chiang, B., (1982). Identifying valid measures of reading. *Exceptional Children*, 49(1), 36-45.
- Deshler, D., Schumaker, J., & Lenz, K. (1984). Academic and cognitive interventions for LD adolescents: Part II. *Annual Review of Learning Disabilities*, 2, 67-76.
- Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology*, 70, 109-118.
- Fuchs, L. S. (1986). Monitoring progress among mildly handicapped pupils: Review of current research and practice. *Remedial and Special Education*, 7, 5-12.
- Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, 9(2), 20-28.
- Germann, G. (1985). *Mainstream consultation procedural manual* (Available from Director of Special Education, Pine County Special Education Cooperative, Sandstone, MN, 55405).
- Good, R. H. III, & Salvia, J. (1988). Curriculum bias in published, norm-referenced reading tests: Demonstrable effects. *School Psychology Review*, 17(1), 51-60.
- Halpern, A. S., & Benz, M. R. (1987). A statewide examination of secondary special education for students with mild disabilities: Implications for the high school curriculum. *Exceptional Children*, 54(2), 122-129.
- Hasbrouck, J. (1987). *Objective and holistic scoring of writing: A manual for collecting and scoring writing samples using holistic and 8 objective scoring procedures*. Unpublished manuscript, University of Oregon.
- Herber, H. (1978). *Teaching reading in content areas* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hinofotis, F. B., & Snow, B. G. (1977). An alternative cloze testing procedure: Multiple-choice format. In J. W. Oller, Jr. & K. Perkins (Eds.) *Research in Language Testing* (pp. 129-133). Rowley, MA: Newbury House.
- Houck, C. K., Geller, C. H., & Engelhard, J. (1988). Learning disabilities teachers' perceptions for adolescents with learning disabilities. *Journal of Learning Disabilities*, 21, 90-97.
- Howard, J. (1983). *Writing to learn*. Washington, DC: Council for Basic Education.
- Howell, K. W., & Kaplan, J. S. (1980). *Diagnosing basic skills*. Columbus, OH: Charles E. Merrill.
- Idol, L. (1986). *Models of curriculum-based assessment*. Rockville, MD: Aspen Systems.
- Jenkins, J. R., & Pany, D. (1978). Standardized achievement tests: How useful for special education? *Exceptional Children*, 44, 448-453.
- Jenkins, J. R., Heliotis, J. D., Stein, M. L., & Haynes, M. C. (1987). Improving reading comprehension by using paragraph restatements. *Exceptional Children*, 54(1), 54-59.
- Kintsch, W., & vanDijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-395.
- Langer, J. (1986). *Children reading and writing*. Norwood, NJ: Ablex.
- Lovitt, T., Rudsit, J., Jenkins, J., Pious, C., & Benedetti, D. (1986). Adapting science materials for regular and learning disabled seventh graders. *Remedial and Special Education*, 7(1), 31-39.
- Marston, D., Lowry, L., Mirkin, P., & Deno, S. (1981). *An analysis of learning trends in simple measures of reading, spelling and written expression: A longitudinal study*. (Research Report No. 49). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- McGee, L. M., & Richgels, D. J. (1985). Teaching expository text structure to elementary students. *The Reading Teacher*, 38, 739-748.
- Meyer, B. J. F. (1977). What is remembered from prose: A function of passage structure. In R. O. Freedle (Ed.), *Discourse production and comprehension* (Vol. 1, pp 307-336). Norwood, NJ: Ablex.
- Niles, O. S. (1985). Integration of Content and Skills Instruction. In T. L. Harris & E. J. Cooper (Eds.), *Reading, thinking, and concept development: Strategies for the classroom*. New York: College Entrance Examination Board.

- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill Book Company.
- Parker, R., Hasbrouck, J., & Tindal, G. (1987). *Valid, objective scoring indices of writing quality: For regular classroom use Grades 2-5*. Unpublished manuscript, University of Oregon.
- Readence, J. E., Bean, T. W., & Baldwin, R. S. (1985). *Content area reading: An integrated approach*. Dubuque, IA: Kendall/Hunt.
- Reynolds, M. C., & Birch, J. W. (1982). *Teaching exceptional children in all America's schools*. Reston, VA: Council for Exceptional Children.
- Roid, G. & Haladyna, T. (1982). *A technology for test-item writing*. New York: Academic Press.
- Schumaker, J., & Deshler, D. (1984). Setting demand variables: A major factor in program planning for LD adolescents. *Topics in Language Disorders Journal*, 4, 22-40.
- Shapiro, E. S., & Derr, T. F. (1987). An examination of overlap between reading curricula and standardized achievement tests. *Journal of Special Education*, 21(2), 59-67.
- Shinn, M. S., Tindal, G., & Stein, S. (1988). Curriculum-based measurement and the identification of mildly handicapped students: A research review. *Professional School Psychology*, 3, 69-85.
- Taylor, B., & Beach, R. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly*, 19(2), 134-146.
- Tierney, R., & Leys, M. (1986). What is the value of connecting reading and writing? In B. Peterson (Ed.), *Convergences: Transactions in reading and writing*. Urbana, IL: National Council of Teachers of English.
- Tierney, R., & Pearson, P. D. (1983). Toward a composing model of reading. *Language Arts*, 60, 568-581.
- Tindal, G., Germann, G., & Deno, S. (1983). *Descriptive research on the Pine County norms: A compilation of findings* (Research Report No. 132). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Tindal, G., & Parker, R. (1989). Assessment of written expression for students in compensatory and special educational programs. *Journal of Special Education*, 23(2), 169-184.
- Tindal, Shinn, M., Walz, L., & Germann, G. (1987). Mainstream consultation in secondary settings: An evaluation of the Pine County model. *Journal of Special Education*, 21(3), 94-106.
- Tucker, J. A. (1986). Curriculum-based assessments: An introduction. *Exceptional Children*, 52, 199-204.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Vacca, R. T. (1981). *Content area reading*. Boston: Little, Brown, & Co.
- Valencia, S. W., & Pearson, P. D. (1988). Principles for classroom comprehension assessment. *Remedial and Special Education*, 9(1), 26-35.
- Videen, J., Deno, S., & Marston, D. (1982). *Correct word sequences: A valid indicator of proficiency in written expression*. Research Report No. 84. Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

APPENDIX A
READING PASSAGES

Greece: Ancient and Modern

The people of Greece are very proud of their history and culture. Greece was a powerful country more than 2000 years ago. Even then it was not a united country, but a loose group of city-states. Its rugged mountains and many islands kept its peoples separated from one another.

Greece has three separate regions: the mainland, the peninsula, and the islands. The northern mainland has wooded hills and fertile land, while to the south is the sea.

Even long ago its closeness to the sea made Greece a seafaring country. Another reason Greece turned to the sea was that its land was poor. And this is still true today. Poor soil, mountains, and a very dry climate make it hard to farm. Greece must buy much of its food from other countries.

Economic development has been slow in Greece. A lack of energy fuels has made it difficult to develop industry. Greece has no natural gas or petroleum deposits. It has no fast-flowing rivers for hydroelectric power. There is some light industry, such as textiles and leather. Commerce also is important, just as it was in ancient times.

Today Greece's chief commercial income comes from shipping. Its shipping fleet is the fifth largest in the world.

After shipping, tourism is Greece's biggest business. Every year thousands of visitors come to Athens to see the ancient ruins. Often they take cruises to Greek islands such as Rhodes or Crete, the home of the ancient Minoan civilization.

The Skeletal and Muscular Systems

Organs working together make up systems. Two of these systems are the skeletal system and the muscular system.

The human skeleton is made up of bone and cartilage. One difference between the two is that cartilage does not contain the calcium or phosphorus compounds that bone contains. This makes cartilage more flexible than bone.

There are 206 bones in the human skeleton. Some of these bones are connected to each other by ligaments. Since ligaments stretch easily, they allow the bones to move freely. This forms what is called a movable joint.

Joints can allow movement in different directions. A hinge joint allows back and forth movement. A ball and socket joint allows rotational movement.

The inside surface of most joints is covered with cartilage. Joints also contain a special fluid that lubricates them so they do not wear each other away.

Movement at the joints and other parts of the body is caused by the muscles. The muscles of the arms and legs are examples of muscles that aid us in movement. These are called voluntary muscles. There are some muscles like the ones found in the digestive, respiratory, and circulatory systems that are involuntary.

All muscles work only by contracting. Since they only work by contracting, they can only pull. They cannot push. If one set of muscles pulls on a tendon to bend a joint, another set of muscles must pull on a different tendon to straighten the same joint.

There are other limiting factors involved with population changes. Higher population densities also cause an increase in disease. Diseased animals become weakened and are easy prey for predators.

Protective Headgear for Motorcycle Operators and Passengers and Moped Riders.

Shall law require motorcycle operators and passengers and moped riders to wear protective headgear?

Motorcycle operator and passenger and moped rider commit offenses if they ride without wearing protective headgear. Motorcycle operator commits offense if passenger not wearing protective headgear.

Measure classifies offenses as Class C traffic infractions, except motorcycle passenger's failure to wear protective headgear, which is not classified by measure. Exempts persons in inclosed cab or operating or riding three-wheeled vehicle at less than 15 miles per hour. Forbids passengers on mopeds.

Under current law, a person who is under 18 years of age commits a Class C traffic infraction if the person operates or rides on a motorcycle or moped without wearing protective headgear.

Also, a person commits a Class C traffic infraction if the person carries on a motorcycle a passenger under 18 years of age who is not wearing protective headgear.

Ballot Measure 2 amends and repeals current law to create four offenses that apply to all persons regardless of age:

- (1) failure of a motorcycle operator to wear protective headgear;
- (2) failure of a motorcycle passenger to wear protective headgear;
- (3) failure of a moped rider to wear protective headgear; and
- (4) endangering a motorcycle passenger.

The last is committed by the operator of a motorcycle if a passenger on the motorcycle is not wearing protective headgear. The measure provides that all the offenses are Class C traffic infractions except for failure of a motorcycle passenger to wear protective headgear, which is an unclassified offense.

APPENDIX B

MAZE TESTS

Student Number _____	Grade _____	School _____
Teacher _____		Date / / 88

Greece: Ancient and Modern

The people of Greece are very proud of their history and culture. Greece was a powerful country (1) than 2000 years ago. Even (2) it was not a united (3), but a loose group of (4) -states. Its rugged mountains and (5) islands kept its peoples separated (6) one another.

Greece has three (7) regions: the mainland, the peninsula, (8) the islands. The northern mainland (9) wooded hills and fertile land, (10) to the south is the (11).

Even long ago its closeness (12) the sea made Greece a (13) country. Another reason Greece turned (14) the sea was that its (15) was poor. And this is (16) true today. Poor soil, mountains, (17) a very dry climate make (18) hard to farm. Greece must (19) much of its food from (20) countries.

Economic development has been (21) in Greece. A lack of (22) fuels has made it difficult (23) develop industry. Greece has no (24) gas or petroleum deposits. It (25) no fast-flowing rivers for (26) power. There is some light (27), such as textiles and leather. (28) also is important, just as (29) was in ancient times.

Today (30) chief commercial income comes from (31). Its shipping fleet is the (32) largest in the world.

After (33), tourism is Greece's biggest business. (34) year thousands of visitors come (35) Athens to see the ancient (36). Often they take cruises to Greek islands such as Rhodes or Crete, the home of the ancient Minoan civilization.

DIRECTIONS:

For each blank in the story, circle the word below that fits the best.

- (1) hydroelectric, buy, more, many, shipping
- (2) and, shipping, buy, natural, then
- (3) fifth, shipping, commerce, many, country
- (4) it, to, every, city, energy
- (5) while, shipping, ruins, sea, many
- (6) from, natural, shipping, sea, slow
- (7) energy, separate, ruins, to, every
- (8) land, and, buy, industry, commerce
- (9) it, fifth, every, has, while
- (10) other, hydroelectric, while, energy, to
- (11) sea, still, it, shipping, separate
- (12) while, from, to, other, fifth
- (13) other, ruins, separate, seafaring, every
- (14) industry, natural, has, to, while
- (15) ruins, many, city, natural, land
- (16) other, to, still, seafaring, shipping
- (17) ruins, commerce, and, country, more
- (18) from, it, energy, ruins, to
- (19) buy, then, and, to, slow
- (20) slow, every, country, has, other
- (21) separate, natural, slow, Greece's, shipping
- (22) energy, fifth, separate, more, then
- (23) other, has, to, many, more
- (24) industry, energy, natural, many, to
- (25) has, and, sea, Greece's, commerce
- (26) commerce, shipping, slow, hydroelectric
- (27) fifth, industry, many, Greece's, ruins
- (28) Many, Buy, Commerce, City, While
- (29) separate, then, many, it, to
- (30) to, commerce, Greece's, shipping, seafaring
- (31) land, from, shipping, many, more
- (32) while, commerce, hydroelectric, fifth, and
- (33) industry, still, and, buy, shipping
- (34) Every, And, Still, Greece's, More
- (35) still, to, buy, then, ruins
- (36) industry, hydroelectric, ruins, energy, buy

Student Number _____ Grade _____ School _____
 Teacher _____ Date ____ / ____ / 88

The Skeletal and Muscular Systems

Organs working together make up systems. Two of these systems are (1) skeletal system and the muscular (2).

The human skeleton is made (3) of bone and cartilage. One (4) between the two is that (5) does not contain the calcium (6) phosphorus compounds that bone contains. (7) makes cartilage more flexible than (8).

There are 206 bones in (9) human skeleton. Some of these (10) are connected to each other (11) ligaments. Since ligaments stretch easily, (12) allow the bones to move (13). This forms what is called (14) movable joint.

Joints can allow (15) in different directions. A hinge (16) allows back and forth movement. (17) ball and socket joint allows (18) movement. The inside surface of (19) joints is covered with cartilage. (20) also contain a special fluid (21) lubricates them so they do (22) wear each other away.

Movement (23) the joints and other parts (24) the body is caused by (25) muscles. The muscles of the (26) and legs are examples of (27) that aid us in movement. (28) are called voluntary muscles. There (29) some muscles like the ones (30) in the digestive, respiratory, (31) circulatory systems that are involuntary.

(32) muscles work only by contracting. (33) they only work by contracting, (34) can only pull. They cannot (35). If one set of muscles pulls on a tendon to bend a joint, another set of muscles must pull on a different tendon to straighten the same joint.

DIRECTIONS:

For each blank in the story, circle the word below that fits the best.

- (1) most, at, muscles, the, since
- (2) difference, joints, system, found, all
- (3) all, up, arms, and, push
- (4) rotational, the, difference, they, by
- (5) they, joints, not, muscles, cartilage
- (6) found, all, they, joint, or
- (7) This, These, They, Bone, Most
- (8) they, freely, bone, the, system
- (9) that, these, the, found, all
- (10) bones, difference, joint, are, cartilage
- (11) or, of, by, they, push, muscles
- (12) movement, rotational, they, push, found
- (13) rotational, joints, they, joint, freely
- (14) that, a, cartilage, movement, bone
- (15) muscles, movement, joint, push, difference
- (16) are, joint, a, they, arms
- (17) Not, Rotational, Muscles, A, Up
- (18) difference, rotational, joint, not, muscles
- (19) freely, this, movement, system, most
- (20) Joints, System, That, At, The
- (21) joints, bones, since, push, that
- (22) and, not, since, all, this
- (23) since, system, freely, joints, of
- (24) found, all, they, cartilage, of
- (25) cartilage, the, bones, joint, arms
- (26) these, arms, bones, rotational, bone
- (27) movement, bones, muscles, cartilage, this
- (28) Are, Movement, These, Difference, Or
- (29) are, not, found, movement, freely
- (30) the, a, movement, and, found
- (31) joints, all, bones, and, rotational
- (32) Not, Found, Muscles, Freely, All
- (33) Freely, They, Since, At, System
- (34) or, this, bones, that, they
- (35) most, rotational, muscles, up, push

APPENDIX C
SCORING SHEETS

Key:

- 1 = partial credit
- 2 = full credit

Student Names/ ID #'s

**Josephine County Secondary Pilot (Spring, 1988)
Content scoring of written retell.**

Scorer: _____ Date Received _____

				1	Greece: Ancient and Modern
					2 The people of Greece are very proud of their history and culture.
					3 Greece was a powerful country more than 2000 years ago.
					4 Even then it was not a united country, but a loose group of city-states.
					5 Its rugged mountains and many islands kept its peoples separated from one another.
					6 Greece has three separate regions:
					7 the mainland, the peninsula, and the islands.
					8 The northern mainland has wooded hills and fertile land,
					9 while to the south is the sea.
					10 Even long ago its closeness to the sea made Greece a seafaring country.
					11 Another reason Greece turned to the sea was that its land was poor.
					12 And this is still true today.
					13 Poor soil, mountains, and a very dry climate make it hard to farm.
					14 Greece must buy much of its food from other countries.
					15 Economic development has been slow in Greece.
					16 A lack of energy fuels has made it difficult to develop industry.
					17 Greece has no natural gas or petroleum deposits.
					18 It has no fast-flowing rivers for hydroelectric power.
					19 There is some light industry, such as textiles and leather.
					20 Commerce also is important,
					21 just as it was in ancient times.
					22 Today Greece's chief commercial income comes from shipping.
					23 Its shipping fleet is the fifth largest in the world.
					24 After shipping, tourism is Greece's biggest business.
					25 Every year thousands of visitors come to Athens to see the ancient
					26 ruins.
					27 Often they take cruises to Greek islands such as Rhodes or Crete,
					A the home of the ancient Minoan civilization.
					B Irrelevant Clauses [use tally marks]
					C Erroneous clauses [use tally marks]
					D Total number of words written (ligible & illegible)
					E #illegible words [write numeral]

Key:

1 = partial credit
2 = full credit

Josephine County Secondary Pilot (Spring, 1988)
Content scoring of written retell.

Scorer: _____	Date Received: _____
---------------	----------------------

Student Names/ ID #'s				
				1 The Skeletal and Muscular Systems
				2 Organs working together make up systems.
				3 Two of these systems are the skeletal system and the muscular system.
				4 The human skeleton is made up of bone and cartilage.
				5 One difference between the two is that cartilage does not contain the calcium or phosphorus compounds that bone contains.
				6 This makes cartilage more flexible than bone.
				7 There are 206 bones in the human skeleton.
				8 Some of these bones are connected to each other by ligaments.
				9 Since ligaments stretch easily,
				10 they allow the bones to move freely.
				11 This forms what is called a movable joint.
				12 Joints can allow movement in different directions.
				13 A hinge joint allows back and forth movement.
				14 A ball and socket joint allows rotational movement.
				15 The inside surface of most joints is covered with cartilage.
				16 Joints also contain a special fluid that lubricates them
				17 so they do not wear each other away.
				18 Movement at the joints and other parts of the body is caused by the muscles.
				19 The muscles of the arms and legs are examples of muscles that aid us in movement.
				20 These are called voluntary muscles.
				21 There are some muscles like the ones found in the digestive, respiratory, and circulatory systems that are involuntary.
				22 All muscles work only by contracting.
				23 Since they only work by contracting, they can only pull.
				24 They cannot push.
				25 If one set of muscles pulls on a tendon to bend a joint,
				26 another set of muscles must pull on a different tendon to straighten the same joint.
				A Irrelevant Clauses [use tally marks]
				B Erroneous clauses [use tally marks]
				C Total number of words written (legible & illegible)
				D # Illegible words [write numeral]
				E Holistic Rating: 1=LOW.....5=HIGH

APPENDIX D
TEACHER DIRECTIONS

Teacher Directions
Administration of Silent Reading & Written Retell

ASSIGNING STUDENT CODE NUMBERS:

Teachers will assign a code number to each student who is attending on the testing day. Students will write these code numbers instead of names on their test papers.

Teachers must save the list of code numbers and names, because each student will use the same code number throughout this Pilot Norming.

Teachers will send the list of code numbers and names to the School District Office at the completion of the Pilot Norming. Only the code numbers (no names) will be sent with test results for analysis to the Univ. of Oregon.

Each teacher should assign code numbers consecutively (1, 2, 3, 4, 5, etc.), in order of seating or class roster. Each student tested by a particular teacher will receive a different code number. If one teacher is testing two or more classrooms, each student in those classrooms should have a unique code number.

TIMING:

The Silent Reading and Written Retell are timed together. From the moment you say "Begin reading", you will write on the board the lapsed time, in 15 second intervals. Students will begin the Written Retell at different times (as soon as each is finished reading). Continue writing the lapsed time on the board through Silent Reading and until the last student has completed his/her Written Retell.

PRELIMINARY:

Write your name and today's date on the board (in the form: Month/Day/Year).
Write the following on the board in large letters and easily visible to all:

TIME:

Have a supply of lined paper on hand. Students will begin writing on the back-side of their passages, but may need more space to write.

DISTRIBUTION AND NUMBERING:

SAY:
"When you get your Written Retell Test please keep the story side down on your desk, do NOT write your name on it."

"I will give you a code number instead of your name to write on the paper, so when the University of Oregon scores the tests, they will not know which paper belongs to which student. Only teachers in this school district will know."

Hand out one Written Retell Test to each student.

SAY:
"Remember to keep the story side down on your desk."
"Write your grade, school, and my name at the top of the sheet. Copy the date from the board. Do it now."
"Now I will give each of you a code number to write at the top of your sheet, where it says, 'Student Number'."

Now read code numbers earlier assigned in order of seating or class roster, or from earlier testing.

SAY:
"John, you are number one, please write it on your test. Mary, you are number two. Fred, you are number three." etc.

Teacher Directions Administration of Maze Test

ASSIGNING STUDENT CODE NUMBERS:

Teachers will assign a code number to each student who is attending on the testing day. Students will write these code numbers instead of names on their test papers.

Teachers must save the list of code numbers and names, because each student will use the same code number throughout this Pilot Norming.

Teachers will send the list of code numbers and names to the School District Office at the completion of the Pilot Norming. Only the code numbers (no names) will be sent with test results for analysis to the Univ. of Oregon.

Each teacher should assign code numbers consecutively (1, 2, 3, 4, 5, etc.), in order of seating or class roster. Each student tested by a particular teacher will receive a different code number. If one teacher is testing two or more classrooms, each student in those classrooms should have a unique code number.

DISTRIBUTION AND NUMBERING:

Write your name and today's date on the board (in the form: Month./Day/Year).

SAY:

"When you get your test please do NOT write your name on it."

Hand out one Maze Test to each student.

SAY:

"We are trying out a kind of reading test, called a Maze. Your score on the test will not count toward your grade. We will study your answers, though, to help plan better ways to teach you reading and to test your reading progress. The test should take less than 15 minutes.

I will give you a code number instead of your name to write on the paper, so when the University of Oregon scores the tests, they will not know which paper belongs to which student. Only teachers in this school district will know."

"Write your grade, school, and my name at the top of the sheet. Copy the date from the board. Do it now."

"Now I will give each of you a code number to write at the top of your sheet, where it says, 'Student Number'."

Now read code numbers earlier assigned in order of seating or class roster.

SAY:

"John, you are number one, please write it on your test. Mary, you are number two. Fred, you are number three." etc.

DIRECTIONS FOR MAZE:

SAY:

"On the left side of the page is a short story, titled: "Greece: Ancient and Modern" (Gr.6) or "The Skeletal and Muscular System" (Gr.8). Some words have been left out of the story. An underlined number is written in the story for each word left out."

"Find the number (1) in the story. Now find the other number (1) on the right hand side of the page, with five words after it. You have to circle the word which you think best fits where the (1) is in the story. Read the story from the beginning, and circle only one of the five words for each number."

"Are there any questions? Start the test now."

FINISHING UP:

The test is not timed.

When students are finished, collect papers and send to Carl Cole at the district office, labeled: **Sec. Pilot Norming**. Early finishers can quietly continue with other seatwork.

DIRECTIONS:**SAY:**

"Please keep the story face down on your desk."

"This is a **Written Retell Test** where you silently read a story, and then write down all you remember about the story. Your score on the test will not count toward your grade. We will study your answers, though, to help plan better ways to teach you reading and to test your reading progress."

"Here is a short story for you to carefully read, titled:

(Gr. 6) "Greece: Ancient and Modern", or

(Gr. 8) "Problems During the Carter Presidency", or

(Gr. 11) "Ending the Slave Trade".

"As soon as you are finished reading, turn over your test sheet, look up at the board, and copy time there into the **DONE READING** box."

"Then, without looking back at the story, on the lines below, write everything you can remember about what you have read. When you are done writing, look up at the board and copy the time into the **DONE WRITING** box."

"Your writing will be judged in two ways: (1) How many facts or ideas from the story you have accurately written down; (2) Your spelling, capitalization/punctuation, sentence structure, and paragraph writing."

"You will get partial credit, even if your spelling, use of vocabulary, sentence structure, or memory of facts is not completely correct."

"You must finish both the reading and writing within 15 minutes. I'll warn you when there are only two minutes left to go."

"If you need more writing space you may come up and get an extra sheet of paper, but only after the sheet in front of you now is full."

"Are there any questions? . . . Begin reading now."

TIMING:

Write on the board the lapsed time, in 15 second intervals. As you write a new lapsed time, erase the old one. The lapsed times should be written like this: :15, :30, :45, 1:00, 1:15, 1:30, 1:45, 2:00 etc.

FINISHING UP:

Even slower students should finish within the fifteen minutes. Early finishers can quietly continue with other seatwork. When all students are finished, collect passages for return to Carl Cole at the district office. Label the return envelope: Sec. Pilot Norming.