Technical Report # 30

District Reading Assessments, Spring 2004 Administration

Julie Alonzo

Gerald Tindal

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

Abstract

This technical report presents an overview of a reading assessment used by a large school district in the Pacific Northwest. It provides sufficient background information to enable those unfamiliar with the assessment to understand the various subtests that comprise it as well as its development and application in the district for making screening, grouping, and evaluation decisions. Compiled performance data from the spring 2004 assessment cycle is presented for all students in the school district in kindergarten through eighth grade for all subtests and for the sample of ninth grade students who were involved in the piloting of the grade nine measures. The technical adequacy of the assessment is explored using both descriptive statistics and item response theory. Suggestions for improvements to the various subtests are discussed.

Introduction

During the 1998 school year the district curriculum and staff development coordinator of a large school district, located in a mid-sized city in the Pacific Northwest, convened a group of 100 kindergarten through third grade teachers who were interested in improving the reading performance of students, as mandated by the school board. They met for a total of 24 hours over the course of the year and by the spring of 1999 had committed to designing and developing a district-wide assessment system to more efficiently and reliably identify K-3 students at risk of reading failure.

That summer, the district curriculum and staff development coordinator formed a task force to begin this work. Members included teachers from both regular education and special education staff, school psychologists, and researchers from the University of Oregon. The group agreed from the start on certain criteria that must be present in any assessment system they might purchase or develop. Each reading kit would be designed to:

1. provide the needed materials to complete individual student reading assessments;
2. use real literature, as much as possible, as opposed to "synthetic" passages developed specifically for the test;
3. assess what teachers thought would be useful in the classroom and yet be psychometrically sound;
4. be brief and easy in its administration;
5. serve as the district's Title One screener;
6. be used in such a way as to show progress over time;
7. provide continuity around testing to assist in the development of district norms; and
8. correlate to the Statewide Test in key benchmark grades.

Initially, the group thought it might find a pre-made assessment kit to purchase for use in the district. Over time, however, as it became evident that none of the reading assessments then available exactly matched their needs, they decided to develop their own.

In the 2001-2002 school year they tested all K-3 students in the district and established district norms. Because grade 4 and 5 teachers began asking for similar assessments at their levels, the district implemented assessments K-5 in the school year 2002-2003. In the spring of 2003 they ran a pilot at the middle level, and in the 2003-2004 school year the district implemented the assessments district-wide and established middle level norms. The district is currently in the process of developing a prototype that can be used at the high school level in grades 9 and 10.

Measurement/Instrument Development

The Task Force identified the critical behaviors of early reading based on a review of the current literature and input from university consultants. The report from the National Reading Panel (2000) was used to focus the discussion on phonemic awareness, phonics, fluency, vocabulary, and comprehension. To address the first two areas identified by the National Reading Panel, the group included phoneme segmentation and letter naming tests for early readers, as the literature clearly identified these skills as predictors of early reading failure. In addition, teachers requested that the district provide assessments of students' abilities to associate particular sounds with letters. Primary teachers also wanted to include a high frequency word protocol to help them determine which high frequency words students actually knew.

The group then moved to consider other factors identified by the National Reading Panel—how to test fluency, vocabulary, and comprehension. Fluency was assessed through a series of Oral Reading Fluency (ORF) measures, using grade-appropriate reading passages. The district decided to

use stories that were not part of the instructional materials being used by district teachers to reduce the chance that student exposure to the passages would inflate the fluency scores. As reading passages were identified, the district decided to use a two-tiered approach to determine reading level. First, they compared the Lexile level of the passage (using a system developed by MetaMetrics) with the Flesch-Kincaid readability index. Second, they selected four passages at each grade level that they thought were equal in reading difficulty and ran a pilot with students to determine their comparability. Once two comparable forms of each MC Comprehension and ORF instrument were developed for each grade level, fourth through eighth, the district designated one form to be used at half the schools in the district in the fall with the other form to be used at the remaining schools. For spring test administrations, the forms are reversed. The earlier grades may eventually follow suit, but currently there is only one form for all students in grades 1 through 3, regardless of school.

Another decision to be made about the tests was whether to include illustrations in the test. Some members of the task force felt strongly that pictures should be included with all primary grade level reading passages (particularly for younger students), because they help readers enhance their construction of meaning. However, the group ultimately decided to include pictures on the reading comprehension measure but not on the oral reading fluency measure because they felt that pictures would distract from the one-minute timing.

Members of the task force thought that for students who had not yet mastered decoding, assessing listening comprehension would be valuable. They decided that first and second grade students whose oral reading fluency was below 39 words per minute would have the comprehension passage (and the comprehension questions – see below) read aloud by the test administrator. Data reports for each student would indicate whether the test was delivered as a listening or reading comprehension measure.

The final comprehension assessment included multiple choice and open-ended items for grades 1 through 9, designed to test literal as well as inferential comprehension. The development of these items occurred over several meetings held during the year in which members of the team openly created question stems and answer options that were then critiqued as a group.

In the primary grades, the District decided to test students on their recognition of high frequency words at the request of teachers (to test vocabulary decoding aimed at structural components of words), but in the intermediate grades they decided to develop traditional vocabulary tests (to test semantic but not syntactic components of words). The group projected that they might find vocabulary to be a better predictor of reading difficulties in older students than phoneme segmentation, high frequency words, and even oral reading rate because understanding vocabulary becomes increasingly important in comprehending text at the upper grade levels. The district used a set of vocabulary lists developed at Ohio State University by Dr. Edgar Dale, a nationally recognized vocabulary expert in the 1980s. Working from these lists, they designed a 60- to 90-item vocabulary measure for each grade to use in the pilot. Following the pilot testing, three forms of a vocabulary test were developed for use at each grade level from third through 8th grade. These forms each contained 25 words and were used during the fall and spring district-wide testing with one additional form available to schools interested in testing in the winter.

To provide a professional format that was clearly accessible to students, the group discussed many particulars related to actually printing the tests including: (a) font and size of type to be used, (b) whether the tests should be timed and, if so, for how long, (c) how scripted the instructions should be, (d) the format of examiner score sheets and student protocols, and (e) the sequencing of tasks, among other issues. The test protocols were formatted to utilize optical character recognition scanning so that all data entry would be automated, and in the fall of 2003, assessments were

administered to all district students in grades K-8, including those receiving special education and English language learner services (Shoemaker & Alonzo, n.d.). Technical reports presenting complete analyses of the results of pilot testing during the 2002-2003 school year can be found at http://www.brtprojects.org/ (Alonzo & Tindal, 2004a, 2004b, 2004c, 2004d; Ketterlin-Geller & Tindal, 2004).

Current Purpose and Uses of the Assessment

Results of all the assessments are entered into a district data-system designed specifically to store the assessment in a secure, on-line format so teachers can access the following: (a) individual student data for use in parent conferences and to inform teaching practices, (b) past assessment data to help teachers understand their students' past reading development and performance on the indicators of reading development, (c) both individual and group data to enable teachers to assess the effectiveness of the curricular and instructional choices they made with their classes, and (d) district data that reflects raw scores, percentile ranks, and district norms for each measure.

In addition to the on-line information, printed student results are distributed to three sets of people: individual teachers receive results related to their own students' performance; school principals receive reports for all students in their school, grouped by teacher; and district administrators receive reports for all students in the district, grouped first by school, then by teacher. Printed report forms are distributed one week after all district testing has been completed.

Individual teachers' use of the data varies greatly. While some use the information to help design their instruction and tailor it to individual student needs, others refer to it only while writing end-of-term report cards. The District Curriculum and Instruction Specialist uses the

information from the assessments to help identify staff development and training needs and to guide her in selecting teachers to receive supplemental assistance. In addition, administrators consider student performance results when evaluating the effectiveness of particular literacy programs being used in the district, to identify staff development needs and to inform their budget plans, as the data may indicate the need to hire or re-assign a reading specialist to meet specific needs.

In addition, a primary use of the K-8 District Reading Assessment results continues to be as a Title I Screener, to identify students in need of intensive intervention. As such, the test is targeted specifically to assess students who struggle with reading with a great deal of accuracy. In the high school grades, however, the district hopes to target students at both ends of the performance spectrum. They would like to develop the assessment to screen students most in need of intensive intervention to make up basic skill deficits as well as to identify students who would most benefit from placement in advanced or honors-level English classes.

Methods

Test Administration

The District Reading Assessment involves a collaborative partnership between the district and a local university. The district provides all the student information to the university, conducts the actual assessments, and distributes the score reports. The university prints the student scan sheets and report forms, analyzes the results, and reports on the psychometric properties of the various test. The District Reading Assessment was administered by a cadre of retired or substitute teachers hired specifically to administer the tests. Test administrators attended an all-day training seminar in which they received instruction on test administration protocols. The

group was organized into teams, with one coordinator per school site whose job included monitoring testers for fidelity of implementation.

For grades K – 1, all tests were administered individually. For the other grades, tests of letter naming, letter sounds, phoneme segmentation, high frequency word lists, sentence reading, oral reading fluency, and listening comprehension were administered individually, and test administrators recorded student scores on a scan sheet. Tests of reading comprehension and vocabulary were group-administered and machine-scored. All test administrators followed standardized written testing protocols.

The Spring District Reading Assessment was given during a four-week testing window immediately following the administration of the statewide large-scale assessment, with the majority of the tests administered in the first three weeks. Each student completed tests on two separate days, with no more than one intervening school day between testing sessions. Table 1 presents an overview of when each of the measures was administered.

Table 1
Overview of When Each Individual Measure was Administered

Measure	Grade Level Administered	Day Administered
Letter naming	K –1	Day 1
Letter sounds	K –1	Day 1
Phoneme segmentation	K – 1	Day 1
High frequency word reading	1 – 2	Day 2
Sentence reading	2	Day 2
Oral reading fluency	1 – 8	Day 1 (1^{st}–2^{nd} grade) Day 2 (3^{rd}–8^{th} grade)
Vocabulary	3 - 8	Day 2
Comprehension (Multiple-choice and Open-ended)	1 – 9	Day 2
Comprehension (maze)	9	Day 2

Description of each measure

Letter naming (Kindergarten)

For the letter naming subtest, students were shown a list of 26 letters of the alphabet, organized randomly in 5 columns on one side of a single sheet of paper. They were instructed to read across each row, saying the name of each letter as they came to it. Test administrators followed along on their own copy of the test protocol, marking any letter the student named incorrectly. Students were given one minute to complete this task. At the end of the minute, test administrators counted the number of letters correctly named and entered the score on the student's scan sheet.

Letter sounds (Kindergarten – 2nd Grade)

The letter sounds sub-test followed a similar protocol. Students were shown a list of 26 letters of the alphabet, organized randomly in 5 columns on one side of a single sheet of paper. They were instructed to read across each row, saying the sound of each letter as they came to it. Test administrators followed along on their own copy of the test protocol, marking any letter the student sounded out incorrectly. To provide teachers with a running record of student mistakes, test administrators wrote down any errors students made during this test. Students were given one minute to complete this task. At the end of the minute, test administrators counted the number of letters correctly sounded out and entered the score on the student's scan sheet.

Phoneme segmentation (Kindergarten)

The phoneme segmentation sub-test required students to correctly segment words on a printed list into their phonemes. The word *dog,* for example, would be read "d – ah – g." The assessors read each word aloud to the student and then asked the student to say the sounds that made up the word. When the assessor read the word *dog,* for example, the student would respond

"d – ah – g." Test administrators marked the phonemes as they were segmented on an individual protocol sheet. To provide teachers with a running record of student mistakes, test administrators wrote down any errors students made during this test. At the end of one minute, test administrators added the number of phonemes correctly segmented and entered the score on the student's scan sheet.

High frequency word reading (1st – 3rd Grade)

The High Frequency Word sub-test consisted of seven potential word lists, organized by level of difficulty (Pre-Primer #1, Pre-Primer #2, Pre-Primer #3, Primer, 1_2, 2_1, and 2_2). Students were administered a series of word lists based on their performance on prior sub-tests. The majority of 1st-Grade students completed the Pre-Primer Word List #3, the Primer Word List, and the 1_2 Word list. First grade students who were unable to complete these more advanced word lists were administered the Pre-Primer Word Lists #1 and #2. The majority of 2nd-Grade students completed the 2_1 and 2_2 Word Lists, but a few lower-performing 2nd graders completed the 1_2 Word List as well. In addition, some Kindergarten students completed both the pre-primer sub-tests.

In all cases, students were presented with a list of high frequency words typed on one side of a single sheet of paper. They were given 40 seconds to read across each row of words before moving down to the next. If students paused for longer than three seconds, they were supplied with the word and told to move on to the next word. Test administrators followed along on their own test protocol, marking any words read incorrectly or skipped. At the end of 40 seconds, test administrators counted the number of words read correctly and entered the score on the student's scan sheet.

Second grade word reading (2nd Grade)

In addition to the High Frequency Word sub-tests, 2nd grade students completed a Word Reading test that included a random sample of phonograms from a list of approximately 2,500 using The Reading Teacher's Book of Lists (Fry, Kress, & Fountoukidis, 2000). Students were presented with 85 single and multi-syllable words typed on one side of a single sheet of paper and given one minute to read as many of them as they could. If students paused for longer than 3 seconds, they were supplied the word and instructed to move on to the next word. Test administrators followed along on their own test protocol, marking any words read incorrectly or skipped. At the end of one minute, test administrators counted the number of words read correctly and entered the score on the student's scan sheet.

Sentence reading (2nd Grade)

A Sentence Reading sub-test also was administered at the second grade. Students were presented with 16 sentences (varying in length from 8 to 19 words) typed on one side of a single sheet of paper and given one minute to read as many of them as possible. The sentences were constructed to reflect a number of structural characteristics of words: prefixes, suffixes, verb tense, syntactic rules relating nouns and verbs as well as conditional clauses and transitions. They were designed to sample students' ability to decode words in context. If students paused for longer than 3 seconds, they were supplied the word and told to move on to the next word. Test administrators followed along on their own test protocol, marking any words read incorrectly or skipped. At the end of one minute, test administrators counted the number of words read correctly and entered the score on the student's scan sheet.

Oral reading fluency (1st – 9th Grade)

The Oral Reading Fluency (ORF) measure was administered individually to each student. Students read aloud for one minute from a grade level passage while test administrators followed along on their own copy of the reading passage, marking any words skipped or read incorrectly. If students paused for longer than 3 seconds, they were supplied the word and instructed to move on to the next word. To provide teachers with a running record of student mistakes, test administrators wrote down any errors students made during this test. At the end of one minute, assessors marked the last word read then counted the total words read, subtracting any words skipped or read incorrectly to arrive at a final ORF score (correct words per minute), which they then entered on the student's scan sheet. Self-corrections and repetitions were counted as correct in determining a student's score.

Vocabulary (3rd – 9th Grade)

The vocabulary measure was group administered to the whole class in a group setting by trained assessors. Students were not given a strict time limit in which to complete this test but most students completed it within approximately 10-15 minutes. Once the majority of the students had finished, test administrators told the remaining students that they would have approximately five more minutes before the test was over. The vocabulary measure was printed on one side of a single sheet of paper and consisted of 25 vocabulary items. The stem consisted of a vocabulary word; below it were three possible answer choices: a correct synonym, a word that reflected a near synonym, and a word not even close to reflecting a synonym. The correct choice (in position of first, second, and third option) was randomly ordered to avoid a pattern effect. Students selected their response by filling in a bubble to the left of the answer choice they believed to be the best definition of those listed for the vocabulary word.

Reading Comprehension (1st – 9th Grade)

The Reading Comprehension sub-test was group administered to students in their regular classroom setting. The test was not timed but generally took about 20-25 minutes. As in the vocabulary test assessors monitored the students, and when the majority had completed the test, students were told they had approximately five more minutes in which to finish. Students read a grade-level appropriate reading passage (See Table 2) and then answered multiple choice (MC) and open-ended (OE) questions about the passage. Both MC and OE questions were written to sample literal and inferential reading comprehension skills. For the MC section of the test, students were provided with four options: one was correct, two were sensible but incorrect, and one was clearly incorrect. The correct choice (in position of first, second, third and fourth option) was randomly ordered to avoid a pattern effect. MC questions were machine scored; items were marked correct if the student selected the appropriate answer choice. All other responses resulted in a score of incorrect. Standardized anchor papers were used by test administrators to score the OE responses. Assessors scored the responses and then entered students' scores on their scan sheets. Table 2 lists the reading passages that were used at each of the grade levels. As mentioned earlier, for grades 4 and above, students in schools from one half the district took Form A while students in the other half took Form B. The forms were reversed in the fall administration.

Table 2
Reading Passages for MC and OE Comprehension Measures

Grade	Title of Story	# of MC Items	# of OE Items	# of Words in Passage
1	*What Will Little Bear Wear?*	5	2	146
2	*Cam Jansen and the Mystery of the Circus Clown*	10	2	228
3*	*The Littles Take a Trip*	25	3	959
4	*Some Writer*	15	2	
	Monkey Business	15	2	906
5	*Bessie Coleman*	15	2	
	The Stag	15	2	672
6	*Shrek*	15	2	
	Beowulf	15	2	1374
7	*Big Things Start Small*	15	2	
	Ali Baba and the Forty Thieves	15	2	1378
8	Swiss Family Robinson	15	2	
	Excerpt from The Life of Honorable Phineas T. Barnum	15	2	1185
9	Excerpt from My Antonia	25	NA	1248

*Grade 3 items were being piloted this year, so the sub-test had 25 items, the 15 best functioning of which will be retained for future use.

Listening Comprehension (1^{st} – 2^{nd} Grade)

In grades 1 and 2, students who scored fewer than 29 words per minute on the ORF sub-test were administered the comprehension sub-test as a *listening* rather than a *reading* comprehension measure. Students listened as test administrators read the comprehension passage aloud. Test administrators then read each question out loud, and the students filled in the bubble on their scan sheet corresponding to their answer choice. Students responded orally to the OE

questions, and test administrators transcribed and then scored their responses using the same standardized written scoring protocol as they used for the reading comprehension sub-tests.

Maze (9th Grade)

In 9th-grade, students were group-administered an un-timed Maze comprehension sub-test, which consisted of a passage from the novel *My Antonia* with approximately every 7th word deleted. Articles, prepositions, conjunctions, and linking verbs were excluded from the every 7th word count in order to select words that would provide information about students' reading comprehension. For each deleted word, students were provided with four multiple choice answer choices: one word correctly completed the sentence both syntactically and semantically, two words were close but incorrect (either syntactically or semantically) in completing the sentence, and one word was clearly incorrect. The correct choice (in position of first, second, third, and fourth option) was randomly ordered to avoid a pattern effect. Students indicated their response choice by filling in a bubble on their scan sheet corresponding to the answer they selected. All responses were machine-scored.

Analysis

Tests of Letter Naming, Letter Sounds, Phonetic Segmenting, High Frequency Words, and Oral Reading Fluency were analyzed descriptively. This analysis was appropriate for these measures because they consisted of a single score rather than a compilation of individual item scores. Student performance on all measures was analyzed using classical statistics to establish district norm scores for the 10th, 25th, 50th, 75th and 90th percentile on each measure at each grade level. In addition, all vocabulary and comprehension measures were analyzed with Item Response Theory (IRT) using a 1 parameter Rasch model in WINSTEPS, v3.48 (Linacre, 2003).

According to the 1PL model (equation 1), the probability of a correct response to a given item *i*, P_i is governed by the person ability parameter (Θ) and the item difficulty parameter (β_i). Given the person's ability, as an item becomes more difficult, the probability of a correct response diminishes. From another perspective, given an item's difficulty, as the person's ability increases, the probability of a correct response increases. According to equation 1, each item has a discrimination parameter. For the 1 PL model, the discrimination parameter (a_i) is typically set to 1.0 uniformly across all items.

Logistic Rasch Model—for item *i*, the Probability of Full Credit as a function of Ability (Θ) and Item Characteristic (β)

$$P_i(\theta) = \frac{e^{a_i(\theta-\beta_i)}}{1+e^{a_i(\theta-\beta_i)}} \text{ (where i = 1, 2, . . . , n)} \qquad \text{(equation 1)}$$

An example control file from the Grade 3 Multiple Choice comprehension test is presented in Figure 1.

Figure 1
Control File for Third-Grade Multiple Choice Reading Comprehension Test

```
*******************************************************************************
*                          * * * *  W I N S T E P S  * * * *
*    -  RASCH ANALYSIS FOR TWO-FACET MODELS  -
*    PERSON, ITEM & RESPONSE STRUCTURE MEASUREMENT AND FIT ANALYSIS
*    INQUIRE:  WINSTEPS
*    PO BOX 811322, CHICAGO  ILLINOIS 60681-1322
*    Tel. & FAX (312) 264-2352
*    www.winsteps.com
*
*    COPYRIGHT (C) JOHN MICHAEL LINACRE, 1991-2003
*    NOVEMBER 26, 2003  VERSION 3.48
*******************************************************************************
 TABLE 0.1 grade 3 Multiple Choice Spring 04       Gr3IRTMC.dat Jun 18 14:10 2004
-------------------------------------------------------------------------------
TITLE= grade 3 Multiple Choice Spring 04
CONTROL FILE: C:\Julie\gr3_sp04_mc.con
OUTPUT FILE: F:\01Julie\IRT Rdg Sp04\Gr3IRTMC.dat
DATE: Jun 18 14:10 2004
0 students     Records Input.

TABLE 0.3 grade 3 Multiple Choice Spring 04       Gr3IRTMC.dat Jun 18 14:10 2004
-------------------------------------------------------------------------------
;grade 3 Multiple Choice Spring 04

&INST
TITLE = "grade 3 Multiple Choice Spring 04"
PERSON = students ; persons are ...
ITEM = item ; items are ...
;DATA = C:\ the path to your data file i.e., DATA=c:\grade3.dat
ITEM1 = 7 ; column 7 is the response to item 1
NI = 25 ; number of items
NAME1 = 1 ; column of first character of person label
NAMELEN = 6 ; length of person label
XWIDE = 1 ; number of columns per item response
CODES = 12349 ; valid codes in data file, (wrong if missing)
;MISSING = 9
KEY1 = 34131241421243334131241 42 ; key for MCQ grade 3 scoring
DISTRACTORS=YES

&END
```

Results

District Norm Scores

The information in Tables 3 through 12 was generated from the tests given in spring 2004. In all cases, only student performance on *required* measures is reported.

Table 3
Kindergarten Data: Spring 2004

	Letter Naming	Segmenting	Letter Sounds
n	1116	1113	1115
M	33.95	29.45	25.50
SD	17.59	16.45	14.05
10th percentile	9	5	6
25th percentile	22	15	14
50th percentile	34	31	26
75th percentile	46	42	36
90th percentile	57	50	42

Table 4
First Grade Data: Spring 2004

	High Freq Pre-Primer 3	High Freq. Primer	High Freq. 1_2	ORF	MC Tot		OE Tot	
					R	L	R	L
n	1197	1156	1147	1196	875	329	876	329
M	18.21	18.53	17.03	70.39	4.7	4.13	3.25	2.97
SD	7.91	7.40	8.21	42.88	0.67	1.05	0.84	1.00
10th percentile	4	5	2	17	4	3	2	2
25th percentile	12	14	11	33.25	5	4	3	2
50th percentile	23	23	21	66.50	5	4	3	3
75th percentile	24	24	23	100	5	5	4	4
90th percentile	24	24	24	133	5	5	4	4

Note: MC and OE total scores are given as scores on either Reading Comprehension (R) or Listening Comprehension (L) tests.

Table 5
Second Grade Data: Spring 2004

	High Freq. 1_2	High Freq. 2_1	High Freq. 2_2	ORF	MC Tot		OE Tot	
					R	L	R	L
n	117	1199	1133	1186	893	307	892	306
M	12.58	22.30	22.18	71.81	7.73	6.36	3.26	2.41
SD	6.90	4.0	4.24	38.96	1.42	2.01	1.09	1.44
10^{th} percentile	3	19	19	20	6	3	2	0
25^{th} percentile	7	23	23	38.75	7	5	2.25	1
50^{th} percentile	13	24	24	71	8	7	4	2
75^{th} percentile	19	24	24	99	9	8	4	4
90^{th} percentile	21	24	24	126.30	9	9	4	4

Note: MC and OE total scores are given as scores on either Reading Comprehension (R) or Listening Comprehension (L) tests.

Table 5, Cont.
Second Grade Data: Spring 2004

	GR2 Word Reading	GR2 Sentence Reading
n	1054	1055
M	35.07	92.31
SD	21.40	44.55
10^{th} percentile	8	35
25^{th} percentile	16	58
50^{th} percentile	34	90
75^{th} percentile	51	120
90^{th} percentile	64	154.4

Table 6
Third Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	1280	1296	1276	1268
M	105.73	22.58	16	3.48
SD	43.09	3.52	4.60	1.72
10^{th} percentile	50	19	9	1
25^{th} percentile	74	22	13	2
50^{th} percentile	105	24	17	4
75^{th} percentile	140	25	19	5
90^{th} percentile	166	25	21	6

Table 7
Fourth Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	1200	1210	1205	1202
M	115.94	21.05	10.85	2.59
SD	35.89	3.94	3.16	1.23
10th percentile	70	16	6	1
25th percentile	91	19	9	2
50th percentile	119	22	12	3
75th percentile	142	24	13	4
90th percentile	162.9	25	14	4

Table 8
Fifth Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	1273	1304	1301	1303
M	135.24	20.94	11.88	2.39
SD	36.55	3.44	2.74	1.40
10th percentile	91.40	16	8	0
25th percentile	109	19	11	1
50th percentile	137	22	12	2
75th percentile	161	23	14	4
90th percentile	181	24	15	4

Table 9
Sixth Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	785	796	795	795
M	141.96	18.89	11.30	2.96
SD	40.32	4.25	2.76	1.11
10th percentile	88.6	13	7	1
25th percentile	117.5	16	10	2
50th percentile	143	19	12	3
75th percentile	169	22	13	4
90th percentile	191	24	14	4

Table 10
Seventh Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	808	827	825	821
M	144.50	16.77	11.25	2.9
SD	38.05	4.32	2.66	1.26
10th percentile	91.9	11	7.6	1
25th percentile	116	14	10	2
50th percentile	146.5	17	12	3
75th percentile	173	20	13	4
90th percentile	193	23	14	4

Table 11
Eighth Grade Data: Spring 2004

	ORF	Voc Tot	MC Tot	OE Tot
n	844	861	860	857
M	137.55	17.44	10.44	3.07
SD	34.20	4.24	3.11	1.14
10th percentile	91	12	6	1
25th percentile	113	15	8	2
50th percentile	138	18	11	4
75th percentile	163	21	13	4
90th percentile	182	23	14	4

Table 12
Ninth Grade Pilot Data: Spring 2004

	ORF	Voc Tot	MC Tot	Maze Tot
n	211	222	212	225
M	151	28.16	13.42	38.39
SD	38.68	4.83	3.78	10.61
10th percentile	101	21	8	22.6
25th percentile	123	25	11	32.0
50th percentile	153	29	14	40.0
75th percentile	178	32	16	46.5
90th percentile	199	34	18	51.0

Item Response Theory Analysis By Measure

Results of the IRT analyses are presented in Tables 13 through 43. Results are reported for each of the multiple-choice formatted measures administered at each grade level. This type of analysis is inappropriate for measures that do not use a multiple choice format, so only multiple choice measures were analyzed using IRT. For each grade, the first table presents item difficulty, standard error, and mean square outfit for each item. This information is used in determining how well an item is functioning in a given test instrument. When using IRT to evaluate item functioning, an item is considered functioning well if its Mean Square Outfit lies between 0.5 and 1.5. Mean Square Outfits less than 0.5 or greater than 1.5 indicate a poorly fitted item that should be revised or removed.

Results of the distractor analysis for each item are presented in the second table for each grade level. When using IRT to evaluate distractor use, distractors are considered functioning well if the Mean Measure of students selecting the correct answer is higher than the Mean Measure of students selecting the other answer choices. In addition, distractors that are selected by no students are considered poorly functioning and in need of revision.

Because there were two forms of the MC Comprehension test for students in grades 4 through 8, each form (designated by title of the reading passage used) is analyzed separately in both item functioning and distractor analyses.

Table 13
Item Functioning: Grade 1 Measures

Item #	Multiple Choice Listening Comprehension			Multiple Choice Reading Comprehension		
	Difficulty	Standard Error	Outfit	Difficulty	Standard Error	Outfit
1	2.28	0.17	1.24	2.5	0.16	1.43
2	-2.09	0.34	1.07	-1.57	0.32	0.76
3	0.71	0.18	0.93	0.27	0.18	1.02
4	-1.05	0.26	0.69	-1.07	0.27	0.83
5	0.16	0.19	0.92	-0.13	0.20	0.98

Table 14
Distractor Analysis for Grade 1, Multiple Choice Comprehension

Item	Answer Choice	MC Listening Comprehension		MC Reading Comprehension	
		# of students who selected this choice	Mean measure of students who selected this choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	57	1.31	75	1.53
	B	16	0.94	14	1.22
	C*	204	3.08	731	3.53
	D	47	0.70	49	1.05
2	A	6	-0.08	4	-1.64
	B	4	-0.46	5	1.10
	C	3	-0.22	6	-1.21
	D*	312	2.42	855	3.24
3	A*	259	2.80	820	3.33
	B	2	0.68	6	-0.72
	C	13	-0.28	4	0.90
	D	51	0.60	39	0.96
4	A	2	-1.29	3	-0.30
	B	3	-0.27	1	-3.37
	C*	300	2.57	849	3.26
	D	20	-0.76	15	0.12
5	A*	275	2.71	831	3.31
	B	22	0.23	15	0.56
	C	4	-2.03	6	0.78
	D	24	0.41	17	0.21

*Indicates correct answer choice.

Table 15
Item Functioning: Grade 2 Measures

Item #	Multiple Choice Listening Comprehension			Multiple Choice Reading Comprehension		
	Difficulty	Standard Error	Outfit MNSQ	Difficulty	Standard Error	Outfit
1	-0.52	0.15	0.86	-0.96	0.12	0.74
2	-0.26	0.14	1.03	-0.25	0.10	0.94
3	-1.05	0.16	0.87	-1.31	0.13	0.97
4	-0.22	0.14	0.81	-0.59	0.11	0.73
5	-0.28	0.14	0.88	-1.06	0.12	0.70
6	3.23	0.20	1.73	4.39	0.12	2.55
7	-1.05	0.16	1.14	-0.26	0.10	0.95
8	0.04	0.14	0.8	-0.49	0.10	1.03
9	0.35	0.13	1.16	1.10	0.08	0.92
10	-0.22	0.14	0.99	-0.59	0.11	0.99

Table 16
Distractor Analysis for Grade 2, Multiple Choice Comprehension

Item	Answer Choice	MC Listening Comprehension		MC Reading Comprehension	
		# of students who selected this choice	Mean measure of students who selected this choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	25	-0.04	35	0.38
	B	12	-0.49	17	0.40
	C*	219	1.11	792	2.05
	D	46	-0.28	37	0.38
2	A	14	-0.07	9	-0.29
	B	60	0.17	138	0.75
	C	21	-0.53	5	0.00
	D*	206	1.09	731	2.13
3	A*	242	1.02	814	1.99
	B	16	-0.78	12	0.26
	C	17	0.40	35	0.87
	D	26	-0.60	21	0.24
4	A	8	-1.10	6	0.16
	B	14	0.03	15	0.12
	C*	204	1.20	763	2.10
	D	72	-0.16	96	0.54
5	A*	207	1.20	799	2.06
	B	11	-0.25	12	0.00
	C	45	-0.19	31	0.15
	D	38	-0.32	39	0.15
6	A	51	0.90	175	1.68
	B*	38	1.48	132	3.21
	C	39	0.41	120	1.48
	D	172	0.61	451	1.68
7	A	13	0.18	27	0.52
	B	29	-0.07	58	0.49
	C	15	0.15	66	0.81
	D*	242	0.92	732	2.14
8	A*	190	1.29	754	2.09
	B	16	0.01	39	0.43
	C	5	-0.33	4	-0.15
	D	89	-0.20	86	0.75
9	A	33	0.05	61	0.83
	B	66	0.37	207	0.93
	C	28	-0.03	58	0.81
	D*	173	1.16	555	2.46
10	A	45	0.39	66	0.78
	B*	204	1.18	763	2.08
	C	22	-0.41	22	0.33
	D	28	-0.85	31	0.49

*Indicates correct answer choice.

Table 17
Item Functioning: Grade 3 Measures

Item #	Vocabulary			Multiple Choice Reading Comprehension		
	Difficulty	Standard Error	Outfit	Difficulty	Standard Error	Outfit
1	1.15	0.09	1.35	-1.77	0.10	0.73
2	-1.25	0.17	0.69	-1.71	0.10	0.61
3	-0.84	0.15	0.88	-1.31	0.09	0.91
4	-0.52	0.14	0.82	-0.73	0.08	0.73
5	1.9	0.08	1.29	-0.83	0.08	0.76
6	-0.96	0.16	0.92	0.37	0.07	1.01
7	0.01	0.12	0.86	-1.04	0.08	0.61
8	0.47	0.10	0.90	-0.84	0.08	0.69
9	-0.35	0.13	0.68	0.49	0.06	1.09
10	0.24	0.11	0.84	-0.67	0.08	0.78
11	-0.16	0.12	0.70	0.38	0.07	1.34
12	-0.3	0.13	1.16	-1.34	0.09	0.72
13	-0.91	0.15	0.36	0.38	0.07	1.11
14	0.51	0.10	0.93	0.64	0.06	1.09
15	-0.96	0.16	0.42	0.38	0.07	0.85
16	-0.48	0.13	0.60	3.98	0.11	4.85
17	-0.76	0.15	0.93	-0.12	0.07	0.98
18	-0.87	0.15	0.65	0.42	0.07	1.04
19	0.97	0.09	1.27	0.30	0.07	0.98
20	2.09	0.08	1.35	2.48	0.07	1.74
21	-0.19	0.12	0.56	0.56	0.06	1.16
22	0.50	0.10	0.98	-0.81	0.08	0.89
23	1.03	0.09	1.21	0.06	0.07	0.95
24	-0.43	0.13	0.46	0.77	0.06	1.12
25	0.09	0.11	0.98	-0.03	0.07	0.79

Table 18
Distractor Analysis for Grade 3 Vocabulary and MC Reading Comprehension Measures

Item	Vocabulary			MC Reading Comprehension		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	1068	3.50	A	61	-0.72
	B	154	2.05	B	24	-0.80
	C	51	1.43	C*	1134	1.10
				D	47	-0.50
2	A	27	0.20	A	46	-0.48
	B*	1246	3.32	B	62	-0.90
	C	13	0.62	C	31	-1.03
				D*	1128	1.12
3	A	24	0.35	A*	1082	1.16
	B*	1230	3.35	B	116	-0.22
	C	27	0.69	C	36	-0.88
				D	27	-1.26
4	A	41	0.78	A	85	-0.72
	B	32	0.65	B	64	-0.05
	C*	1214	3.38	C*	998	1.26
				D	114	-0.26
5	A*	960	3.66	A*	1014	1.24
	B	61	1.48	B	32	-0.98
	C	264	2.08	C	129	-0.29
				D	84	-0.27
6	A	16	0.70	A	167	-0.19
	B	24	-0.12	B*	775	1.37
	C*	1235	3.34	C	40	-0.35
				D	278	0.56
7	A	64	1.07	A	101	-0.20
	B*	1180	3.43	B	54	-0.77
	C	41	1.07	C	51	-0.89
				D*	1045	1.23
8	A	72	1.05	A*	1015	1.25
	B*	1142	3.49	B	113	-0.33
	C	64	1.16	C	46	-0.18
				D	84	-0.51
9	A*	1204	3.40	A	139	0.11
	B	38	0.80	B	300	0.72
	C	45	0.58	C	66	-0.79
				D*	747	1.34
10	A	55	0.90	A	105	-0.19
	B*	1162	3.47	B*	987	1.26
	C	61	0.98	C	51	-0.54
				D	112	-0.15
11	A*	1192	3.43	A*	772	1.17
	B	56	0.76	B	136	0.29
	C	27	0.29	C	96	0.49
				D	252	0.77
12	A*	1201	3.39	A	39	-0.76
	B	29	0.68	B*	1086	1.16
	C	54	1.34	C	47	-0.57
				D	80	-0.25

13	A*	1233	3.38	A	44	-0.62
	B	28	-0.19	B	334	0.57
	C	21	0.04	C	97	-0.10
				D*	771	1.33
14	A	106	1.69	A	51	-0.20
	B	39	0.08	B	323	0.49
	C*	1138	3.50	C*	710	1.37
				D	165	0.36
15	A	23	0.25	A	277	0.36
	B*	1235	3.36	B	109	0.01
	C	21	0.06	C*	772	1.44
				D	95	-0.36
16	A	28	0.57	A	314	0.72
	B	33	0.32	B	272	0.97
	C*	1212	3.39	C*	94	0.59
				D	563	1.13
17	A*	1226	3.36	A	94	-0.08
	B	42	0.77	B	208	0.32
	C	15	-0.07	C	55	-0.41
				D*	883	1.30
18	A*	1231	3.36	A*	763	1.39
	B	30	0.39	B	216	0.24
	C	25	0.25	C	150	0.35
				D	109	0.27
19	A	59	1.52	A	68	-0.22
	B	130	1.81	B	313	0.44
	C*	1089	3.51	C*	790	1.36
				D	69	-0.06
20	A	268	2.28	A*	287	1.46
	B*	930	3.70	B	217	0.30
	C	73	1.19	C	38	-0.50
				D	694	1.05
21	A	31	0.18	A	434	0.62
	B*	1194	3.44	B*	729	1.32
	C	48	0.60	C	41	-0.61
				D	34	-0.67
22	A	52	1.32	A	73	-0.42
	B*	1139	3.48	B	47	-0.49
	C	87	1.27	C	95	0.28
				D*	1011	1.21
23	A	97	1.75	A*	844	1.34
	B*	1082	3.52	B	174	0.42
	C	96	1.77	C	44	-0.53
				D	166	0.12
24	A*	1209	3.42	A	43	-0.26
	B	45	0.42	B	81	-0.36
	C	28	0.06	C	425	0.73
				D*	679	1.36
25	A	49	0.83	A	85	-0.05
	B*	1174	3.43	B*	864	1.39
	C	53	1.44	C	137	-0.03
				D	146	-0.01

*Indicates correct answer choice.

Table 19
Item Functioning: Grade 4 Vocabulary Measure

Item #	Difficulty	Standard Error	Outfit
1	-1.78	0.17	0.93
2	-1.57	0.16	0.98
3	-0.58	0.11	0.64
4	-0.85	0.12	0.73
5	0.11	0.09	1.06
6	0.49	0.08	0.89
7	-0.34	0.10	0.77
8	0.84	0.08	0.88
9	0.14	0.09	1.13
10	0.61	0.08	1.10
11	-0.77	0.12	1.06
12	-0.32	0.10	0.68
13	-0.66	0.11	0.82
14	-0.98	0.13	0.56
15	0.04	0.09	0.75
16	1.25	0.07	1.09
17	-0.02	0.09	0.91
18	0.36	0.09	1.13
19	1.84	0.07	1.11
20	0.95	0.08	1.17
21	-0.68	0.11	0.86
22	1.57	0.07	1.23
23	-0.59	0.11	1.00
24	1.60	0.07	1.29
25	-0.66	0.11	0.82

Table 20
Distractor Analysis for Grade 4 Vocabulary Measure

Item #	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	1161	2.55
	B	14	0.01
	C	17	0.38
2	A	10	0.26
	B*	1153	2.57
	C	34	0.48
3	A	64	0.64
	B*	1093	2.68
	C	37	0.58
4	A	27	0.36
	B	55	0.70
	C*	1114	2.64
5	A*	1025	2.73
	B	31	0.22
	C	138	1.37
6	A	99	0.69
	B	112	1.28
	C*	975	2.84
7	A	41	0.57
	B*	1072	2.71
	C	82	0.82
8	A	224	1.25
	B*	923	2.90
	C	39	0.52
9	A*	1021	2.73
	B	89	1.12
	C	84	1.18
10	A	49	0.76
	B*	958	2.79
	C	179	1.48
11	A*	1108	2.64
	B	68	0.92
	C	14	-0.18
12	A*	1070	2.72
	B	60	0.59
	C	56	0.65
13	A*	1100	2.66
	B	59	0.96
	C	33	0.11
14	A	29	0.19
	B	34	0.32
	C*	1122	2.64
15	A	91	0.98
	B*	1033	2.76
	C	57	0.65
16	A	242	1.58
	B	82	1.00
	C*	854	2.94
17	A*	1040	2.74
	B	54	0.46
	C	91	1.12
18	A*	993	2.77
	B	114	1.39
	C	72	0.91
19	A	100	1.12
	B	340	1.60
	C*	742	3.12
20	A	233	1.54
	B*	905	2.84
	C	43	0.93
21	A	46	0.99
	B*	1101	2.64
	C	39	0.62
22	A	379	1.66
	B*	794	2.96
	C	19	-0.18
23	A	26	0.57
	B*	1094	2.65
	C	61	0.85
24	A*	788	2.95
	B	62	0.52
	C	343	1.84
25	A	44	0.59
	B*	1100	2.66
	C	40	0.52

*Indicates correct answer choice.

Table 22
Item Functioning: Grade 4 MC Comprehension Measures

Item #	Story 1 (*Some Writer)*			Story 2 (*Monkey Business)*			
	Difficulty	Standard Error	Outfit	Item #	Difficulty	Standard Error	Outfit
3	-1.58	0.15	1.27	8	-1.27	0.15	0.56
12	-1.11	0.13	0.81	11	-1.12	0.14	0.61
13	-1.00	0.13	0.63	5	-0.90	0.13	0.77
8[2]	-0.92	0.13	0.46	10	-0.81	0.13	0.91
9[2]	-0.92	0.13	0.80	6	-0.59	0.12	0.73
11[2]	-0.73	0.12	0.77	1[2]	-0.22	0.11	1.30
7[2]	-0.72	0.12	0.77	2[2]	-0.21	0.11	0.91
4[2]	-0.70	0.12	1.05	7	-0.10	0.11	0.74
10	-0.27	0.11	0.85	13	0.04	0.11	0.94
15	0.10	0.11	0.71	12	0.47	0.10	1.10
5	0.29	0.10	1.16	9	0.77	0.10	0.95
2	0.41	0.10	0.99	3[2]	0.84	0.10	1.06
6	0.81	0.10	1.06	4[2]	0.86	0.10	1.42
14	1.54	0.10	1.34	14	0.93	0.10	1.20
1[1]	4.80	0.17	9.90	15	1.31	0.10	1.02

[1]Indicates item is misfitting.
[2]Indicates item is redundant.

Table 23
Distractor Analysis for Grade 4 MC Reading Comprehension Measures

Item	Story 1 (*Some Writer)*			Story 2 (*Monkey Business)*		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	59	-0.24	A	37	0.16
	B	23	-0.25	B	35	0
	C*	42	0.99	C*	472	1.81
	D	475	1.65	D	57	0.84
2	A	105	0.56	A	59	0.33
	B	62	-0.52	B	17	-0.25
	C	36	0.66	C	55	0.03
	D*	396	1.90	D*	471	1.90
3	A*	535	1.59	A*	377	2.08
	B	29	-0.41	B	73	0.50
	C	21	-1.17	C	58	0.67
	D	15	-0.61	D	92	0.48
4	A	29	-0.91	A	21	-0.54
	B	26	-0.68	B	22	0.23
	C*	486	1.72	C*	375	1.93
	D	57	0.30	D	183	1.05
5	A*	407	1.80	A*	517	1.81
	B	81	0.71	B	14	0
	C	15	-1.33	C	36	0
	D	90	0.36	D	35	-0.61
6	A	140	0.45	A	43	0.16
	B*	357	1.99	B*	498	1.86
	C	71	0.53	C	27	-0.30
	D	26	-0.17	D	33	-0.44
7	A	47	-0.01	A	61	0.32
	B	33	-0.35	B	51	-0.30
	C	28	-1.04	C	25	-0.23
	D*	487	1.73	D*	463	1.98
8	A*	500	1.75	A*	536	1.78
	B	33	-1.08	B	23	-0.60
	C	27	-0.67	C	23	-0.45
	D	36	-0.38	D	19	-0.88
9	A	37	-0.73	A	88	0.49
	B	43	0.13	B	89	0.18
	C	15	-1.32	C	35	0.67
	D*	500	1.70	D*	384	2.15
10	A	22	-0.55	A	31	-0.31
	B*	455	1.79	B*	512	1.79
	C	19	-0.48	C	38	0.43
	D	97	0.21	D	17	-0.31
11	A*	488	1.72	A*	529	1.80
	B	57	-0.05	B	28	-0.52
	C	16	-1.32	C	20	-0.40
	D	34	-0.17	D	18	-0.75
12	A	21	-0.56	A	34	0.19
	B*	511	1.65	B*	413	1.98
	C	52	-0.33	C	14	-0.20
	D	12	-0.23	D	137	0.74
13	A	31	-0.42	A	108	0.30
	B	22	-0.72	B	15	-0.18
	C	35	-0.61	C	19	0.07
	D*	505	1.70	D*	451	1.95
14	A	152	1.07	A	124	0.82
	B	75	0.32	B	76	0.63
	C*	282	2.11	C*	368	2.07
	D	85	0.29	D	28	0.41
15	A	51	-0.35	A	105	1.02
	B	35	-0.55	B	118	0.67
	C*	424	1.95	C*	329	2.24
	D	85	0.28	D	43	-0.07

*Indicates correct answer choice.

Table 24
Item Functioning: Grade 5 Vocabulary Measure

Item #	Difficulty	Standard Error	Outfit
1	-1.18	0.06	1.00
2	-4.86	0.14	1.43
3	2.03	0.15	1.65
4	1.12	0.10	1.08
5	1.15	0.10	1.24
6	2.13	0.16	1.14
7	-2.99	0.07	1.05
8	2.29	0.17	2.10
9	3.12	0.25	1.88
10	2.67	0.20	1.65
11	-5.43	0.18	1.88
12	-3.20	0.08	1.03
13	2.80	0.22	1.03
14	2.44	0.18	1.42
15	1.11	0.10	1.07
16	3.42	0.29	1.18
17	-2.34	0.06	1.01
18	1.69	0.13	1.08
19	2.71	0.21	1.34
20	-3.83	0.09	1.45
21	-3.44	0.08	1.11
22	-2.88	0.07	1.16
23	0.69	0.09	1.17
24	0.78	0.09	1.13
25*	NA	NA	NA

*Item 25 was inadvertently left off the test protocols.

Table 25
Distractor Analysis for Grade 5 Vocabulary Measure

Item #	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	558	-1.23
	B	638	-1.71
	C	102	-1.81
2	A	11	-1.81
	B*	1246	-1.49
	C	42	-2.04
3	A	157	-1.68
	B*	46	-1.19
	C	1097	-1.50
4	A	992	-1.54
	B	197	-1.55
	C*	106	-1.08
5	A*	103	-1.08
	B	583	-1.53
	C	598	-1.56
6	A	43	-1.68
	B	1210	-1.52
	C*	42	-1.02
7	A	74	-2.11
	B*	1031	-1.40
	C	141	-1.84
8	A	1237	-1.51
	B*	36	-1.30
	C	25	-1.77
9	A*	16	-1.38
	B	1228	-1.51
	C	52	-1.40
10	A	1232	-1.51
	B*	25	-0.88
	C	33	-1.79
11	A*	1269	-1.50
	B	20	-2.13
	C	12	-1.58
12	A*	1070	-1.40
	B	56	-2.00
	C	172	-2.03
13	A*	22	-0.79
	B	6	-1.08
	C	1272	-1.52
14	A	25	-1.44
	B	1239	-1.52
	C*	31	-1.16
15	A	19	-1.50
	B*	107	-0.99
	C	1163	-1.55
16	A	1221	-1.52
	B	63	-1.47
	C*	12	-0.88
17	A*	884	-1.33
	B	62	-1.85
	C	348	-1.89
18	A*	63	-1.06
	B	27	-1.40
	C	1203	-1.53
19	A	36	-2.07
	B	1221	-1.50
	C*	24	-0.71
20	A	93	-1.78
	B*	1162	-1.47
	C	30	-2.01
21	A	83	-1.95
	B*	1110	-1.42
	C	99	-2.09
22	A	244	-1.87
	B*	1009	-1.40
	C	42	-1.93
23	A	998	-1.53
	B*	153	-1.14
	C	139	-1.75
24	A*	142	-1.10
	B	1072	-1.55
	C	78	-1.73
25**	A	0	0
	B*	0	0
	C	0	0

*Indicates correct answer choice.
**Item 25 was inadvertently left off the test protocols.

Table 26
Item Functioning: Grade 5 MC Comprehension Measures

	Story 1 (*Bessie Coleman)*				Story 2 (*The Stag)*		
Item #	Difficulty	Standard Error	Outfit	Item #	Difficulty	Standard Error	Outfit
6	-1.45	0.17	0.73	6	-1.38	0.16	0.67
1[2]	-0.84	0.14	0.95	7	-0.68	0.13	0.93
10[2]	-0.84	0.14	1.17	11	-0.48	0.12	1.09
4	-0.63	0.13	0.87	9	-0.40	0.12	0.82
14	-0.49	0.13	0.75	2	-0.37	0.12	0.79
2	-0.32	0.12	0.84	14[2]	-0.09	0.11	0.84
11	-0.17	0.12	0.65	15[2]	-0.09	0.11	0.79
9	-0.12	0.12	0.88	5	0.00	0.11	1.16
13	0.15	0.11	0.91	8	0.03	0.11	0.97
3	0.32	0.11	1.06	4	0.05	0.11	0.92
12	0.43	0.11	0.97	10	0.08	0.11	1.12
7	0.66	0.10	1.25	1	0.23	0.10	1.00
15	0.69	0.10	1.10	3	0.63	0.10	1.06
5[2]	1.31	0.10	1.03	12	1.10	0.09	1.06
8[2]	1.31	0.10	1.20	13	1.37	0.09	1.13

[1]Indicates item is misfitting.
[2]Indicates item is redundant.

Table 27
Distractor Analysis for Grade 5 MC Reading Comprehension Measures

Item	Story 1 (*Bessie Coleman)*			Story 2 (*The Stag)*		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	12	0.42	A	25	0.34
	B	17	0.06	B	95	0.97
	C*	532	2.01	C*	545	2.18
	D	39	0.45	D	35	0.64
2	A	31	0.36	A	53	0.37
	B	8	0.03	B	34	0.44
	C	57	0.56	C	17	0.43
	D*	502	2.08	D*	595	2.13
3	A*	455	2.17	A*	505	2.24
	B	55	0.61	B	67	0.66
	C	36	0.62	C	93	1.22
	D	54	0.96	D	32	0.53
4	A	44	0.41	A	107	0.71
	B	10	-0.19	B	14	0.37
	C*	521	2.06	C*	561	2.18
	D	25	0.20	D	18	0.40
5	A*	361	2.38	A*	566	2.11
	B	51	0.11	B	17	0.70
	C	162	1.36	C	36	0.68
	D	25	0.33	D	81	1.04
6	A	11	0.07	A	13	0.17
	B*	557	1.97	B*	651	2.02
	C	7	0.15	C	13	0.33
	D	23	-0.22	D	22	-0.33
7	A	78	1.24	A	35	0.56
	B	36	1.08	B	29	0.23
	C	61	0.75	C	20	0.53
	D*	425	2.15	D*	616	2.08
8	A*	361	2.31	A*	563	2.16
	B	45	0.82	B	13	0.82
	C	13	0.14	C	72	0.53
	D	180	1.24	D	52	0.94
9	A	23	-0.18	A	24	0.50
	B	12	0.12	B	25	0.15
	C	74	0.75	C	52	0.46
	D*	489	2.12	D*	597	2.14
10	A	23	0.51	A	72	1.00
	B*	532	2.00	B*	559	2.15
	C	21	0.32	C	26	0.60
	D	24	0.42	D	43	0.56
11	A*	492	2.17	A*	603	2.07
	B	32	0.10	B	19	0.11
	C	42	0.22	C	60	1.04
	D	30	0.35	D	17	0.18
12	A	56	0.66	A	153	1.34
	B*	446	2.21	B*	453	2.37
	C	73	0.93	C	51	0.31
	D	22	0.07	D	41	0.50
13	A	17	-0.14	A	14	-0.07
	B	40	0.57	B	134	1.34
	C	70	0.78	C	128	1.01
	D*	469	2.17	D*	421	2.39
14	A	10	-0.39	A	35	0.20
	B	66	0.39	B	45	0.40
	C*	513	2.09	C*	573	2.19
	D	8	0.13	D	46	0.81
15	A	36	0.88	A	23	0.52
	B	96	0.94	B	27	0.17
	C*	422	2.22	C*	573	2.19
	D	44	0.75	D	74	0.56

*Indicates correct answer choice.

Table 28
Item Functioning: Grade 6 Vocabulary Measures

Item #	Difficulty	Standard Error	Outfit
1	-1.66	0.11	0.58
2	0.83	0.06	1.23
3	-0.08	0.07	0.88
4	0.34	0.06	0.98
5	0.16	0.07	1.03
6	-1.09	0.09	0.74
7	-0.45	0.07	0.70
8	-0.08	0.07	0.84
9	0.33	0.06	0.99
10	0.26	0.07	1.01
11	-0.70	0.08	0.87
12	-0.06	0.07	0.82
13	-0.33	0.07	0.85
14	0.76	0.06	1.60
15	-0.18	0.07	0.97
16	-0.43	0.07	0.94
17	0.18	0.07	0.82
18	0.58	0.06	0.94
19	0.10	0.07	0.86
20	-0.37	0.07	1.30
21	0.42	0.06	1.00
22	0.74	0.06	0.96
23	0.65	0.06	1.19
24	-0.30	0.07	0.97
25	0.37	0.06	0.92

Table 29
Distractor Analysis for Grade 6 Vocabulary Measure

Item	Vocabulary		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	1279	1.51
	B	84	-0.05
	C	13	-0.51
2	A	471	1.05
	B*	832	1.70
	C	74	0.18
3	A	269	0.62
	B*	1051	1.67
	C	52	-0.08
4	A	51	0.53
	B	369	0.68
	C*	955	1.72
5	A*	997	1.67
	B	133	0.54
	C	234	0.78
6	A	73	0.29
	B	84	0.11
	C*	1219	1.55
7	A	206	0.30
	B*	1123	1.66
	C	45	-0.02
8	A	203	0.55
	B*	1049	1.70
	C	119	0.19
9	A*	958	1.71
	B	183	0.64
	C	229	0.69
10	A	101	0.57
	B*	975	1.70
	C	295	0.70
11	A*	1165	1.57
	B	134	0.52
	C	74	0.16
12	A*	1045	1.69
	B	173	0.49
	C	153	0.40
13	A*	1100	1.64
	B	59	-0.15
	C	216	0.57
14	A	14	-0.46
	B	514	1.27
	C*	849	1.50
15	A	189	0.65
	B*	1071	1.64
	C	114	0.38
16	A	127	0.54
	B	126	0.51
	C*	1118	1.60
17	A*	993	1.76
	B	357	0.52
	C	25	-0.34
18	A*	896	1.78
	B	248	0.59
	C	219	0.78
19	A	264	0.64
	B	98	0.28
	C*	1010	1.71
20	A	154	1.00
	B*	1107	1.57
	C	106	0.27
21	A	412	0.73
	B*	935	1.72
	C	26	0.20
22	A	143	0.49
	B*	854	1.82
	C	373	0.79
23	A	336	0.94
	B*	877	1.68
	C	157	0.78
24	A*	1094	1.63
	B	149	0.51
	C	132	0.47
25	A	141	0.43
	B*	949	1.74
	C	279	0.72

*Indicates correct answer choice.

Table 30
Item Functioning: Grade 6 MC Comprehension Measures

Item #	Story 1 (*Beowulf)*			Story 2 (*The Making of Shrek)*			
	Difficulty	Standard Error	Outfit	Item #	Difficulty	Standard Error	Outfit
13[1]	-2.60	0.24	0.32	9	-1.55	0.14	0.57
11[2]	-1.08	0.14	0.66	10[1]	-1.21	0.12	0.49
14[2]	-1.08	0.14	0.90	13	-0.66	0.11	0.82
10	-0.89	0.13	0.60	14	-0.61	0.11	1.04
12	-0.46	0.12	0.68	4	-0.41	0.10	1.26
3	-0.31	0.11	1.21	12	-0.35	0.10	0.88
6	-0.24	0.11	0.80	8	-0.05	0.10	1.10
15[2]	0.12	0.1	0.91	3	0.02	0.10	1.12
5[2]	0.13	0.1	0.85	15	0.19	0.09	0.94
9[2]	0.14	0.1	0.96	6	0.26	0.09	1.15
8	0.18	0.1	0.81	7	0.55	0.09	1.03
1	0.65	0.1	0.99	1	0.74	0.09	1.12
2	0.69	0.1	1.14	11	0.84	0.09	1.03
7	1.21	0.09	0.93	5	1.09	0.09	1.23
4[1]	3.56	0.11	3.90	2	1.16	0.09	0.93

[1]Indicates item is misfitting.
[2]Indicates item is redundant.

Table 31
Distractor Analysis for Grade 6 MC Reading Comprehension Measures

Item #	Story 1 (*Beowulf*)			Story 2 (*The Making of Shrek*)		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	24	0.33	A	199	0.45
	B	80	0.75	B	85	0.72
	C*	471	2.12	C*	433	1.78
	D	97	1.00	D	2	2.08
2	A	22	0.73	A	34	-0.01
	B	115	1.16	B	267	0.53
	C	71	0.86	C	38	0.54
	D*	467	2.06	D*	379	2.01
3	A*	561	1.95	A*	520	1.66
	B	26	0.44	B	18	0.76
	C	30	0.97	C	53	-0.03
	D	59	0.59	D	128	0.37
4	A	494	1.77	A	63	0.66
	B	8	0.52	B	12	-1.26
	C*	136	1.94	C*	564	1.54
	D	37	0.69	D	80	0.39
5	A*	524	2.07	A*	388	1.82
	B	9	-0.09	B	213	0.75
	C	134	0.64	C	54	0.28
	D	9	0.39	D	61	0.68
6	A	37	0.46	A	108	0.62
	B*	555	2.02	B*	493	1.67
	C	27	0.40	C	80	0.38
	D	57	0.40	D	33	-0.05
7	A	80	0.68	A	164	0.67
	B	108	1.05	B	54	0.24
	C	80	0.99	C	42	-0.23
	D*	406	2.28	D*	458	1.78
8	A*	519	2.09	A*	528	1.65
	B	27	-0.05	B	65	-0.45
	C	82	0.75	C	105	0.85
	D	48	0.54	D	18	-0.44
9	A	84	0.74	A	27	-0.65
	B	16	0.19	B	23	-1.15
	C	50	0.70	C	22	-0.50
	D*	523	2.05	D*	647	1.51
10	A	30	0.01	A	21	-0.77
	B*	599	1.96	B*	627	1.57
	C	21	-0.07	C	45	-0.82
	D	25	-0.01	D	25	-0.30
11	A*	609	1.92	A*	421	1.83
	B	35	0.53	B	34	-0.35
	C	19	-0.46	C	177	0.79
	D	9	-0.41	D	86	0.33
12	A	14	0.49	A	79	0.02
	B*	572	2.01	B*	558	1.65
	C	40	0.03	C	13	-0.84
	D	50	0.26	D	64	0.22
13	A	7	-0.52	A	80	-0.32
	B	4	-1.23	B	19	-0.21
	C	7	-1.23	C	32	0.29
	D*	657	1.81	D*	586	1.61
14	A	16	-0.11	A	43	-0.29
	B	24	-0.34	B	27	-0.32
	C*	609	1.92	C*	582	1.58
	D	25	0.50	D	65	0.34
15	A	80	0.69	A	5	-0.26
	B	37	0.76	B	181	0.38
	C*	525	2.06	C*	501	1.74
	D	33	0.11	D	32	-0.42

*Indicates correct answer choice.

Table 32
Item Functioning: Grade 7 Vocabulary Measure

Item #	Difficulty	Standard Error	Outfit
1	-0.51	0.07	1.26
2	-0.23	0.06	1.19
3	-0.33	0.06	1.22
4	0.16	0.06	1.02
5	-0.76	0.07	1.04
6	-0.15	0.06	0.95
7	-0.89	0.07	0.82
8	1.10	0.06	0.98
9	0.22	0.06	1.09
10	-0.15	0.06	0.95
11	-1.00	0.07	0.83
12	0.69	0.06	1.10
13	-0.27	0.06	1.03
14	0.47	0.06	0.86
15	0.72	0.06	0.89
16	-0.07	0.06	1.26
17	-0.21	0.06	0.87
18	0.12	0.06	0.87
19	0.75	0.06	0.96
20	-0.35	0.06	1.06
21	0.00	0.06	0.93
22	0.38	0.06	1.00
23	1.13	0.06	1.08
24	-0.48	0.07	1.06
25	-0.36	0.07	0.81

Table 33
Distractor Analysis for Grade 7 Vocabulary Measure

Item #	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	1015	0.98
	B	226	0.58
	C	110	0.04
2	A	286	0.55
	B*	950	1.01
	C	108	0.09
3	A	338	0.50
	B*	974	0.98
	C	41	0.18
4	A	132	-0.03
	B	367	0.51
	C*	849	1.12
5	A*	1069	1.00
	B	89	-0.04
	C	192	0.32
6	A	266	0.29
	B	146	0.21
	C*	929	1.09
7	A	128	0.01
	B*	1094	1.02
	C	129	0.09
8	A	449	0.34
	B*	590	1.37
	C	302	0.55
9	A*	832	1.10
	B	60	-0.21
	C	457	0.50
10	A	143	0.19
	B*	931	1.10
	C	251	0.30
11	A*	1115	1.00
	B	114	0.02
	C	116	0.09
12	A*	703	1.19
	B	425	0.46
	C	220	0.44
13	A*	959	1.06
	B	304	0.40
	C	84	-0.10
14	A	296	0.28
	B	288	0.24
	C*	763	1.29
15	A	516	0.38
	B*	695	1.34
	C	131	0.05
16	A	302	0.48
	B	130	0.32
	C*	909	1.03
17	A*	946	1.12
	B	281	0.25
	C	118	0.01
18	A*	859	1.20
	B	186	0.22
	C	299	0.21
19	A	237	0.31
	B	419	0.41
	C*	686	1.29
20	A	234	0.31
	B*	979	1.04
	C	117	0.25
21	A	289	0.25
	B*	891	1.15
	C	167	0.18
22	A	223	0.27
	B*	788	1.19
	C	333	0.39
23	A	740	0.51
	B*	582	1.31
	C	23	-0.30
24	A*	1009	1.02
	B	156	0.23
	C	180	0.36
25	A	184	0.02
	B*	981	1.12
	C	178	0.15

*Indicates correct answer choice.

Table 34
Item Functioning: Grade 7 MC Comprehension Measures

Item #	Story 1 (*Big Things Start Small)*			Item #	Story 2 (*Ali Baba)*		
	Difficulty	Standard Error	Outfit		Difficulty	Standard Error	Outfit
9	-1.15	0.13	0.50	10	-1.58	0.16	0.57
3	-1.01	0.13	0.75	5	-1.50	0.16	0.64
8	-0.85	0.12	0.61	3	-1.43	0.15	1.12
10	-0.71	0.12	0.75	8	-1.29	0.15	0.66
13	-0.70	0.12	0.68	9	-1.25	0.14	0.68
14	-0.61	0.12	0.90	2	-0.83	0.13	1.19
15	-0.45	0.11	0.98	1	-0.51	0.12	1.25
5	-0.40	0.11	0.82	4	-0.26	0.11	0.88
4	-0.39	0.11	0.85	14	-0.10	0.10	0.83
2	0.04	0.10	1.24	6	1.00	0.09	0.90
6	0.66	0.09	1.04	12	1.08	0.09	1.06
11	0.79	0.09	1.39	13	1.32	0.09	0.98
12	0.98	0.09	1.14	7	1.70	0.09	1.34
1	1.65	0.09	1.05	11	1.80	0.09	1.34
7	2.16	0.09	1.24	15	1.85	0.09	1.00

[1]Indicates item is misfitting.
[2]Indicates item is redundant.

Table 35
Distractor Analysis for Grade 7 MC Reading Comprehension Measures

Item #	Story 1 (*Big Things Start Small)*			Story 2 (*Ali Baba)*		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	112	0.94	A	54	1.06
	B	18	0.29	B	593	1.77
	C*	320	2.21	C*	9	-0.17
	D	218	0.83	D	39	0.84
2	A	59	0.57	A	615	1.77
	B	86	0.91	B	62	0.80
	C	19	-0.34	C	16	0.31
	D*	506	1.77	D*	4	-0.11
3	A*	587	1.72	A*	13	-0.45
	B	58	0.16	B	15	0.42
	C	8	-0.76	C	24	0.81
	D	17	-0.39	D	646	1.74
4	A	77	0.42	A	108	0.69
	B	18	-0.66	B	573	1.87
	C*	544	1.80	C*	8	-0.89
	D	33	0.13	D	7	0.29
5	A*	545	1.79	A*	10	-0.21
	B	59	0.30	B	5	-1.42
	C	15	-0.47	C	33	0.23
	D	53	0.30	D	649	1.76
6	A	48	-0.13	A	39	0.69
	B*	441	1.94	B*	29	0.42
	C	133	1.05	C	435	2.10
	D	47	0.34	D	193	0.97
7	A	248	1.03	A	341	2.01
	B	76	1.06	B	35	0.88
	C	89	0.97	C	240	1.42
	D*	258	2.25	D*	81	1.00
8	A*	577	1.78	A*	30	0.20
	B	35	-0.09	B	19	-0.16
	C	22	-0.67	C	8	0.46
	D	37	-0.14	D	640	1.78
9	A	24	-0.74	A	638	1.78
	B	37	-0.17	B	28	0.10
	C	15	-0.81	C	21	-0.05
	D*	595	1.74	D*	11	0.43
10	A	29	-0.47	A	9	-0.19
	B*	568	1.80	B*	20	0.09
	C	36	0.31	C	652	1.75
	D	38	-0.44	D	16	-0.17
11	A*	426	1.81	A*	169	1.31
	B	49	0.50	B	177	1.35
	C	89	0.91	C	23	0.42
	D	106	1.19	D	327	2.05
12	A	32	-0.43	A	6	0.69
	B*	403	1.96	B*	238	1.15
	C	182	0.92	C	424	2.04
	D	53	1.11	D	29	-0.04
13	A	37	-0.15	A	393	2.15
	B	18	-0.64	B	246	1.06
	C	45	0.04	C	21	0.00
	D*	567	1.79	D*	36	0.96
14	A	24	-0.54	A	8	-0.11
	B	33	0.27	B	559	1.90
	C*	561	1.75	C*	39	0.34
	D	50	0.52	D	87	0.68
15	A	53	0.26	A	26	0.70
	B	56	0.37	B	321	2.30
	C*	549	1.76	C*	36	0.69
	D	11	0.04	D	311	1.15

*Indicates correct answer choice.

Table 36
Item Functioning: Grade 8 Vocabulary Measure

Item #	Difficulty	Standard Error	Outfit
1	0.20	0.06	1.02
2	-0.71	0.07	0.93
3	0.08	0.06	0.90
4	0.04	0.06	0.82
5	0.22	0.06	1.28
6	-0.33	0.07	0.92
7	0.65	0.06	1.14
8	0.04	0.06	0.85
9	0.01	0.06	1.29
10	1.01	0.06	1.27
11	-1.01	0.08	0.70
12	0.89	0.06	0.95
13	0.08	0.06	0.87
14	0.14	0.06	1.34
15	-0.22	0.07	0.87
16	0.03	0.06	1.01
17	-0.62	0.07	0.73
18	-0.68	0.07	0.71
19	-0.28	0.07	0.86
20	-0.53	0.07	0.93
21	0.50	0.06	1.20
22	0.18	0.06	1.00
23	-0.10	0.06	0.89
24	0.50	0.06	0.91
25	-0.10	0.06	0.89

Table 37

Distractor Analysis for Grade 8 Vocabulary Measure

Item #	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	915	1.28
	B	307	0.50
	C	162	0.48
2	A	103	0.20
	B*	1127	1.17
	C	161	0.41
3	A	117	0.25
	B*	949	1.30
	C	323	0.45
4	A	255	0.37
	B	177	0.27
	C*	957	1.32
5	A*	911	1.18
	B	88	0.37
	C	375	0.75
6	A	165	0.41
	B	176	0.33
	C*	1048	1.22
7	A	341	0.85
	B*	793	1.27
	C	252	0.45
8	A	206	0.19
	B*	958	1.32
	C	221	0.47
9	A*	966	1.16
	B	379	0.80
	C	42	-0.45
10	A	598	0.88
	B*	689	1.25
	C	93	0.15
11	A*	1180	1.19
	B	47	-0.14
	C	162	0.08
12	A*	723	1.45
	B	426	0.66
	C	235	0.32
13	A*	948	1.32
	B	236	0.28
	C	199	0.45
14	A	336	0.82
	B	121	0.38
	C*	931	1.17
15	A	206	0.37
	B*	1022	1.26
	C	153	0.23
16	A	58	0.15
	B	367	0.54
	C*	961	1.24
17	A*	1109	1.24
	B	118	0.06
	C	161	0.20
18	A*	1120	1.24
	B	91	0.06
	C	172	0.10
19	A	95	0.19
	B	252	0.35
	C*	1037	1.26
20	A	173	0.35
	B*	1090	1.20
	C	110	0.22
21	A	52	0.01
	B*	833	1.22
	C	503	0.77
22	A	323	0.59
	B*	921	1.28
	C	145	0.25
23	A	203	0.33
	B*	993	1.28
	C	191	0.37
24	A*	834	1.38
	B	405	0.51
	C	145	0.29
25	A	308	0.44
	B*	992	1.27
	C	86	0.07

*Indicates correct answer choice.

Table 38
Item Functioning: Grade 8 MC Comprehension Measures

Item #	Story 1 (*P.T. Barnum)*			Item #	Story 2 (*Swiss Family Robinson)*		
	Difficulty	Standard Error	Outfit		Difficulty	Standard Error	Outfit
3	-1.21	0.12	0.78	3	-1.41	0.13	0.82
1	-0.67	0.10	0.73	10	-1.10	0.12	0.66
2	-0.66	0.10	0.70	5	-0.75	0.11	0.89
9	-0.45	0.10	1.12	9	-0.47	0.10	0.80
8	-0.35	0.10	1.01	6	-0.46	0.10	0.91
5	-0.26	0.10	0.86	4	-0.16	0.10	0.78
10	0.18	0.09	0.88	11[2]	0.03	0.10	0.87
6[2]	0.19	0.09	1.04	12[2]	0.03	0.10	0.91
15[2]	0.19	0.09	0.88	8	0.05	0.09	0.76
12	0.31	0.09	1.05	1	0.11	0.09	1.00
13	0.36	0.09	0.90	7	0.57	0.09	1.21
14	0.45	0.09	1.16	2	0.60	0.09	1.29
4	0.46	0.09	1.45	15	0.64	0.09	1.18
7	0.69	0.09	1.13	13	1.15	0.09	1.23
11	0.75	0.09	0.95	14	1.17	0.09	1.20

[1]Indicates item is misfitting.
[2]Indicates item is redundant.

Table 39
Distractor Analysis for Grade 8 MC Reading Comprehension Measures

Item	Story 1 (*P.T. Barnum)*			Story 2 (*Swiss Family Robinson)*		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	48	-0.33	A	42	-0.13
	B	38	-0.12	B	25	-0.54
	C*	555	1.37	C*	484	1.64
	D	66	-0.20	D	143	0.53
2	A	44	-0.63	A	141	0.73
	B	80	-0.27	B	30	-0.29
	C	28	0.13	C	97	0.69
	D*	554	1.39	D*	426	1.62
3	A*	601	1.27	A*	616	1.43
	B	22	-0.08	B	15	-0.70
	C	63	-0.43	C	35	-0.18
	D	21	-0.33	D	30	-0.35
4	A	53	-0.19	A	85	0.02
	B	211	0.82	B	35	-0.09
	C*	426	1.34	C*	514	1.66
	D	15	-0.55	D	60	0.00
5	A*	513	1.43	A*	569	1.52
	B	60	0.13	B	51	-0.11
	C	118	-0.06	C	45	0.06
	D	17	-0.63	D	29	-0.38
6	A	98	0.14	A	15	-0.47
	B*	460	1.48	B*	544	1.55
	C	69	-0.09	C	52	0.02
	D	81	0.46	D	81	0.25
7	A	84	0.00	A	29	0.10
	B	196	0.59	B	169	0.62
	C	31	-0.05	C	63	0.41
	D*	397	1.54	D*	430	1.67
8	A*	523	1.37	A*	491	1.72
	B	23	-0.29	B	70	-0.12
	C	71	0.49	C	46	-0.05
	D	91	-0.20	D	88	0.24
9	A	81	0.36	A	31	-0.81
	B	34	-0.05	B	49	-0.28
	C	58	-0.26	C	70	0.30
	D*	533	1.33	D*	545	1.60
10	A	132	0.15	A	32	-0.65
	B*	461	1.53	B*	596	1.49
	C	53	-0.10	C	44	-0.10
	D	60	0.11	D	23	-0.49
11	A*	389	1.67	A*	494	1.68
	B	73	0.22	B	89	0.10
	C	82	0.19	C	56	0.04
	D	162	0.28	D	53	0.28
12	A	62	0.10	A	43	0.15
	B*	445	1.48	B*	493	1.67
	C	123	0.25	C	109	0.06
	D	74	0.35	D	48	0.33
13	A	26	-0.74	A	247	0.86
	B	76	-0.13	B	37	-0.17
	C	164	0.38	C	48	0.22
	D*	439	1.57	D*	357	1.78
14	A	164	0.58	A	31	-0.14
	B	58	0.01	B	208	0.68
	C*	427	1.47	C*	354	1.87
	D	56	-0.03	D	99	0.56
15	A	103	0.06	A	152	0.76
	B	64	0.10	B	49	0.21
	C*	460	1.52	C*	421	1.70
	D	81	0.14	D	71	0.12

*Indicates correct answer choice.

Table 40
Grade 9 Vocabulary and MC Reading Comprehension Measures

Item #	Vocabulary			Multiple Choice Reading Comprehension		
	Difficulty	Standard Error	Outfit	Difficulty	Standard Error	Outfit
1	0.33	0.18	0.75	-0.20	0.15	1.06
2	0.29	0.18	0.67	0.87	0.15	1.07
3	-1.00	0.26	1.32	0.71	0.15	1.21
4	-0.46	0.22	0.86	-0.77	0.16	0.87
5	-2.07	0.39	0.77	-1.92	0.21	0.77
6	-0.32	0.21	1.54	0.36	0.15	0.94
7	1.01	0.16	1.12	-0.50	0.15	0.86
8	-0.76	0.24	1.03	1.26	0.16	1.25
9	1.19	0.16	0.79	0.36	0.15	0.98
10	-0.76	0.24	1.93	1.08	0.16	1.18
11	-0.32	0.21	0.87	-1.75	0.20	0.79
12	0.45	0.17	1.09	-0.65	0.16	0.89
13	1.35	0.15	1.31	0.80	0.15	1.22
14	-0.82	0.24	0.79	2.28	0.21	1.31
15	-1.92	0.37	0.58	0.26	0.15	1.06
16	-0.80	0.24	0.77	0.80	0.15	1.04
17	0.59	0.17	0.66	-0.41	0.15	1.15
18	0.59	0.17	1.20	-0.50	0.15	0.83
19	1.09	0.16	1.44	-1.32	0.18	0.73
20	-0.24	0.20	0.70	0.19	0.15	1.09
21	-0.41	0.21	0.76	-1.92	0.21	0.67
22	2.85	0.17	1.59	0.58	0.15	0.92
23	-1.47	0.31	0.52	0.43	0.15	0.86
24	-1.00	0.26	0.62	1.06	0.16	1.61
25	0.65	0.17	1.01	-1.10	0.17	0.70
26	-0.32	0.21	0.59			
27	0.13	0.18	0.81			
28	-0.20	0.20	0.83			
29	-0.04	0.19	1.00			
30	-1.38	0.3	0.66			
31	-1.68	0.33	0.63			
32	1.84	0.15	1.11			
33	-0.76	0.24	0.67			
34	2.53	0.16	0.98			
35	-0.08	0.19	0.97			
36	1.93	0.15	1.12			

Table 41
Distractor Analysis for Grade 9 Vocabulary and MC Reading Comprehension Measures

Item	Vocabulary			MC Reading Comprehension		
	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A*	167	2.11	A	15	-0.69
	B	29	0.84	B	53	0.24
	C	20	1.07	C*	126	0.39
				D	16	-0.54
2	A	17	0.9	A	98	0.12
	B*	168	2.15	B	31	-0.05
	C	31	0.72	C	3	-1.03
				D*	77	0.50
3	A	4	0.93	A*	84	0.36
	B*	198	1.9	B	38	-0.20
	C	14	1.28	C	10	-0.94
				D	79	0.40
4	A	22	1.09	A	24	-0.55
	B	6	0.61	B	17	-0.72
	C*	188	1.97	C*	150	0.47
				D	19	-0.08
5	A*	209	1.9	A*	185	0.34
	B	3	0.9	B	8	-1.11
	C	4	-0.33	C	13	-0.49
				D	7	-0.44
6	A	21	1.26	A	82	-0.03
	B	10	1.09	B*	100	0.59
	C*	185	1.95	C	26	-0.38
				D	5	-0.50
7	A	41	1.38	A	13	-0.72
	B*	142	2.11	B	22	-0.37
	C	28	1.29	C	37	-0.07
				D*	139	0.49
8	A	14	1.02	A*	61	0.50
	B*	194	1.94	B	93	0.28
	C	8	1	C	33	-0.14
				D	24	-0.25
9	A*	135	2.31	A	49	-0.03
	B	4	0.35	B	39	-0.26
	C	77	1.11	C	24	0.04
				D*	100	0.56
10	A	10	0.32	A	21	-0.36
	B*	194	1.93	B*	68	0.49
	C	11	1.76	C	118	0.20
				D	5	-1.16
11	A*	185	1.99	A*	181	0.36
	B	9	0.51	B	12	-0.85
	C	22	1.16	C	11	-0.55
				D	7	-0.90
12	A*	163	2.07	A	8	-0.89
	B	22	1.04	B*	145	0.45
	C	28	1.42	C	40	-0.20
				D	20	-0.32
13	A*	128	2.08	A	51	0.35
	B	42	1.34	B	46	0.19
	C	44	1.73	C	35	-0.33
				D*	80	0.37
14	A	15	0.94	A	41	-0.09
	B	3	-0.18	B	10	-0.83
	C*	195	1.96	C*	29	0.42
				D	2	-0.73
15	A	4	0.47	A	30	-0.23
	B*	208	1.9	B	62	0.02
	C	4	0.52	C*	105	0.44
				D	15	0.26
16	A	15	0.94	A	8	-0.97
	B	6	0.88	B	9	-0.79
	C*	194	1.95	C*	80	0.57
				D	115	0.13
17	A*	158	2.22	A	62	0.21
	B	37	0.99	B	10	-0.76

	C	21	0.52	C	6	-1.09
				D*	135	0.34
18	A*	158	2.02	A*	139	0.51
	B	29	1.41	B	38	-0.21
	C	28	1.34	C	22	-0.37
				D	13	-0.67
19	A	46	1.68	A	19	-0.61
	B	31	1.17	B	12	-0.42
	C*	139	2.05	C*	169	0.43
				D	10	-0.83
20	A	13	0.55	A*	108	0.42
	B*	183	2.05	B	9	-0.93
	C	20	0.82	C	86	0.14
				D	8	-0.45
21	A	8	0.96	A	4	-1.00
	B*	187	2	B*	185	0.35
	C	21	0.77	C	15	-0.78
				D	7	-0.37
22	A	63	1.65	A	77	0.07
	B*	66	2.26	B	35	-0.22
	C	86	1.67	C	8	-0.75
				D*	90	0.60
23	A	7	0.17	A*	97	0.65
	B*	204	1.93	B	70	0.07
	C	5	0.7	C	37	-0.47
				D	4	-0.75
24	A*	198	1.95	A	44	-0.23
	B	12	0.56	B	4	-0.80
	C	6	0.77	C	90	0.58
				D*	69	0.15
25	A	40	0.97	A	15	-0.78
	B*	156	2.12	B*	162	0.47
	C	20	1.45	C	20	-0.54
				D	11	-0.48
26	A*	185	2.05			
	B	20	0.81			
	C	11	0.23			
27	A	38	0.95			
	B*	173	2.06			
	C	4	0.96			
28	A	29	1.02			
	B*	182	2.02			
	C	5	0.21			
29	A	20	1.36			
	B	18	0.94			
	C*	178	1.99			
30	A*	203	1.94			
	B	3	0.07			
	C	10	0.48			
31	A	2	1.51			
	B	7	0.51			
	C*	206	1.91			
32	A	86	1.53			
	B*	107	2.29			
	C	21	0.94			
33	A	18	0.68			
	B*	194	1.98			
	C	4	0.43			
34	A*	78	2.54			
	B	99	1.42			
	C	36	1.58			
35	A	9	0.51			
	B*	179	2.01			
	C	25	1.14			
36	A*	103	2.26			
	B	57	1.54			
	C	53	1.39			

Table 42
Grade 9 MAZE Measure

Item	Difficulty	Standard Error	Outfit	Item	Difficulty	Standard Error	Outfit
1	-2.05	0.27	0.94	31	-1.21	0.20	0.83
2	-2.21	0.28	1.22	32	-0.59	0.17	1.02
3	-0.56	0.17	1.18	33	0.94	0.15	1.56
4	-1.21	0.20	1.24	34	0.36	0.15	0.88
5	-0.91	0.19	1.73	35	-0.01	0.16	1.02
6	0.72	0.15	0.98	36	0.83	0.15	1.03
7	-1.57	0.23	0.64	37	-0.98	0.19	1.02
8	0.38	0.15	1.13	38	2.55	0.18	1.17
9	1.48	0.15	1.05	39	-0.09	0.16	0.86
10	-0.37	0.16	1.13	40	-0.56	0.17	0.79
11	0.57	0.15	0.81	41	-0.78	0.18	0.65
12	-0.91	0.19	0.96	42	2.90	0.20	1.84
13	1.15	0.15	0.85	43	0.51	0.15	0.81
14	-1.57	0.23	0.73	44	0.94	0.15	0.89
15	-0.88	0.18	1.01	45	0.60	0.15	0.93
16	2.52	0.18	1.53	46	0.33	0.15	1.04
17	2.86	0.19	1.59	47	1.96	0.16	0.95
18	1.53	0.15	1.59	48	-1.09	0.19	1.25
19	0.77	0.15	1.12	49	-1.13	0.20	1.14
20	-0.98	0.19	1.00	50	-0.81	0.18	0.90
21	-0.21	0.16	0.99	51	-0.48	0.17	0.91
22	2.17	0.17	1.30	52	0.31	0.15	0.92
23	-1.57	0.23	0.93	53	-1.29	0.21	0.99
24	-1.63	0.23	0.59	54	-0.98	0.19	1.15
25	1.91	0.16	1.35	55	-0.16	0.16	0.77
26	-0.65	0.17	1.24	56	-0.09	0.16	0.83
27	1.17	0.15	0.77	57	-0.51	0.17	1.24
28	0.04	0.15	0.79	58	-1.13	0.20	0.78
29	-2.05	0.27	0.94	59	-0.06	0.16	0.87
30	-2.21	0.28	1.22	60	-0.40	0.17	0.83

Table 43
Distractor Analysis for Grade 9 Maze Measure

Item	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice	Item	Answer Choice	# of students who selected this choice	Mean measure of students who selected this choice
1	A	10	0.18	31	A	20	0.26
	B	2	0.78		B	83	1.08
	C*	209	0.94		C*	110	0.91
	D	3	-0.03		D	11	0.46
2	A	4	-0.23	32	A	57	0.38
	B	2	1.32		B	21	0.47
	C	7	-0.24		C	9	-0.39
	D*	211	0.94		D*	137	1.24
3	A*	174	1.01	33	A*	153	1.11
	B	46	0.59		B	21	0.11
	C	3	-1.08		C	10	0.67
	D	1	-0.06		D	38	0.56
4	A	21	0.55	34	A	7	-0.06
	B	6	0.48		B	68	0.60
	C*	193	0.97		C*	115	1.23
	D	4	-0.54		D	34	0.55
5	A*	185	0.88	35	A*	187	1.04
	B	32	1.09		B	20	0.16
	C	3	-0.47		C	5	-0.75
	D	4	0.88		D	12	0.43
6	A	85	0.58	36	A	21	0.28
	B*	120	1.26		B*	44	1.53
	C	16	-0.07		C	7	0.40
	D	4	0.04		D	151	0.80
7	A	10	-0.22	37	A	49	0.27
	B	3	-0.32		B	5	-0.36
	C	10	-0.19		C	14	0.41
	D*	201	1.01		D*	156	1.17
8	A*	136	1.12	38	A*	174	1.12
	B	7	-0.04		B	11	-0.07
	C	75	0.70		C	6	-0.39
	D	7	-0.68		D	34	0.20
9	A	79	0.90	39	A	24	-0.14
	B	21	-0.28		B	12	-0.44
	C	40	0.51		C	7	0.13
	D*	85	1.34		D*	181	1.14
10	A	31	0.61	40	A	69	0.65
	B*	167	1.05		B*	34	1.01
	C	9	-0.20		C	51	0.82
	D	17	0.21		D	69	1.12
11	A*	127	1.34	41	A*	130	1.32
	B	52	0.31		B*	21	0.04
	C	40	0.37		C	41	0.28
	D	5	-0.13		D	32	0.44
12	A	3	-0.97	42	A	89	0.50
	B*	185	1.05		B*	110	1.37
	C	22	0.50		C	7	-0.23
	D	14	-0.14		D	19	0.35
13	A	68	0.38	43	A	26	0.14
	B	19	0.53		B	58	0.59
	C	36	0.42		C	15	0.14
	D*	100	1.46		D*	126	1.27
14	A	8	-0.49	44	A	19	0.12
	B	11	-0.06		B	27	1.15
	C*	201	1.01		C*	138	1.16
	D	5	0.03		D	39	0.14
15	A	8	-0.22	45	A	52	0.36
	B	1	-1.33		B	22	0.46
	C*	184	1.02		C*	65	1.57
	D	32	0.43		D	84	0.78
16	A	6	-0.44	46	A	15	0.18
	B	23	0.81		B	6	-0.35
	C*	45	1.28		C*	190	1.01

	D	151	0.83		D	13	0.52
17	A	100	0.91	47	A	10	-0.06
	B	79	0.75		B	14	0.43
	C	11	0.36		C	10	0.12
	D*	35	1.28		D*	191	1.01
18	A*	83	0.95	48	A*	182	1.06
	B	99	0.96		B	31	0.24
	C	40	0.62		C	7	0.01
	D	2	0.43		D	5	-0.13
19	A	26	0.52	49	A	20	0.16
	B	57	0.79		B	28	0.44
	C*	118	1.17		C*	171	1.10
	D	24	0.11		D	5	-0.76
20	A*	187	1.02	50	A*	139	1.22
	B	20	0.27		B	33	0.35
	C	11	0.64		C	20	0.40
	D	6	-0.63		D	32	0.38
21	A	24	0.65	51	A	23	0.01
	B*	161	1.11		B*	195	1.01
	C	13	-0.10		C	5	0.66
	D	26	0.19		D	2	-0.66
22	A	30	0.50	52	A	23	0.41
	B	56	0.97		B	7	-0.05
	C	80	0.75		C	6	-0.41
	D*	57	1.20		D*	187	1.03
23	A*	200	0.98	53	A*	159	1.19
	B	10	0.25		B	9	0.25
	C	10	0.27		C	31	0.06
	D	3	-0.56		D	26	0.22
24	A	17	-0.19	54	A	12	0.24
	B	4	-0.34		B	12	-0.36
	C	2	-0.95		C	43	0.22
	D*	202	1.02		D*	156	1.21
25	A	58	0.79	55	A	17	0.27
	B*	67	1.23		B*	172	1.03
	C	78	0.64		C	8	0.18
	D	22	0.96		D	26	0.48
26	A*	177	1.01	56	A*	191	1.04
	B	20	0.46		B	13	-0.11
	C	15	0.19		C	9	0.31
	D	13	0.67		D	7	-0.48
27	A	71	0.43	57	A	46	0.20
	B*	99	1.51		B*	155	1.18
	C	35	0.37		C	8	-0.04
	D	19	0.31		D	10	0.48
28	A	27	-0.15	58	A	13	0.06
	B	6	0.26		B	22	0.02
	C	41	0.39		C	14	0.41
	D*	151	1.23		D*	168	1.14
29	A	19	-0.02	59	A	10	0.07
	B	5	0.52		B	36	0.41
	C*	193	1.03		C*	153	1.21
	D	7	-0.01		D	15	-0.13
30	A	9	-0.09	60	A	41	0.42
	B	37	0.50		B	23	0.13
	C*	175	1.04		C*	145	1.24
	D	4	0.06		D	3	-0.50

Discussion

Because each grade level was administered its own distinct set of measures, the results of both the descriptive and IRT analyses are presented for each successive grade. More general conclusions and recommendations applicable to a number of grade levels are presented at the end of this section. The paper concludes with a summary of the changes being made in the District Reading Assessment for the 2004 – 2005 assessment cycle.

Kindergarten Measures

All three of the kindergarten measures appear to be functioning well. They discriminate reliably between students of different skill levels. No changes in the kindergarten measures are recommended at this time.

Grade 1 Measures

Descriptive statistics indicate that the High Frequency Word lists, ORF, and OE Comprehension measures are functioning well. They discriminate reliably between students of different skill levels. No changes in these first-grade measures are recommended at this time; however, there may be a ceiling effect for the MC Comprehension measure. Currently, it appears to be too easy to discriminate reliably between students of different skill levels. IRT analyses indicate that the items are functioning appropriately, but because most of the students are answering all five items correctly, this particular test may not be providing useful information about student comprehension skill. Additional, more difficult, items should be added to the measure or the existing items should be revised to make them more challenging.

Grade 2 Measures

The ORF, Grade 2 Word Reading, and Grade 2 Sentence Reading measures appear to be functioning well according to the descriptive statistics. They discriminate reliably between

students of different skill levels. No changes in these measures are recommended at this time; however, there may be a ceiling effect for the High Frequency Word lists and OE Comprehension measures. Currently, they may be too easy to discriminate reliably between students of different skill levels. The OE Comprehension items appear to be functioning appropriately, but because most of the students are answering both items correctly, this particular test may not provide useful information about student comprehension skill, especially for students taking the OE Reading Comprehension test. The OE Listening Comprehension test appears to be slightly better at differentiating students of different skill levels. Additional and more difficult items should be added to the measure, or the existing items should be revised to make them more challenging. Furthermore, IRT analyses indicate that Item #6 of the MC Comprehension measure is poorly fitted for both the Reading (Mean Square Outfit = 2.55) and the Listening (Mean Square Outfit = 1.73). This item should be revised or replaced.

Grade 3 Measures

The ORF and both MC and OE Comprehension measures appear to be functioning well. They discriminate reliably between students of different skill levels. No changes in the ORF measure are recommended at this time. However, the district would like to abbreviate the comprehension measure to 15 MC and 2 OE items to align it with the measures used in the upper grades. IRT analyses indicate that two MC Comprehension items are poorly fitted and should be dropped: Item #16 (Mean Square Outfit = 4.85) and Item #20 (Mean Square Outfit = 1.74). In addition, Item #s 2, 8, 11, 12, 13, 15, 20, and 22 are redundant—they have the same or very similar difficulty as other items on the test—and can also be dropped.

There also appears to be a ceiling effect for the Vocabulary measure: It may be too easy to discriminate reliably between students of different skill levels. The vocabulary items are

functioning appropriately based on IRT analysis, but because most of the students are answering almost all items correctly, this particular test is not providing useful information about the range of student skill. Additional and more difficult items should be added to the measure, or the existing items should be revised to make them more challenging.

Grade 4 Measures

ORF and open-ended comprehension.

The ORF and OE Comprehension measures are functioning well. They discriminate reliably between students of different skill levels. No changes in the ORF or OE Comprehension measures are recommended at this time.

Multiple choice comprehension passages.

Some Writer.

Item #1 of the MC Comprehension measure is poorly fitted (Mean Square Outfit = 4.80), and an analysis of distractors indicates that the highest performing students are incorrectly selecting answer choice “D” rather than the correct answer “C”. This item should be re-written. In addition, four items are redundant with one another, having difficulty levels within .01 of one another. Items 8 and 9 are roughly equivalent, as are items 7 and 11. This redundancy reduces the information provided by the test to some degree. We recommend either re-writing one of the items in each item pair in an attempt to differentiate the two better or replacing one of the items in each item pair with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

Monkey Business.

Items 1 and 2 are redundant with one another, having difficulty levels within .01 of one another. This redundancy reduces the information provided by the test to some degree. As

before, we recommend either re-writing one of the items in each item pair in an attempt to differentiate the two better or replacing one of the items in each item pair with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

Vocabulary

Again, there appears to be a ceiling effect for the Vocabulary measure. As in the other grade levels, it is too easy to discriminate reliably between students of different skill levels. The vocabulary items function appropriately based on IRT analysis, but because most of the students are answering almost all items correctly, this particular test is not providing useful information about the range of student skill. Additional and more difficult items should be added to the measure, or the existing items should be revised to make them more challenging.

Grade 5 Measures

ORF and open-ended comprehension.

The ORF and OE Comprehension measures are functioning well. They discriminate reliably between students of different skill levels. No changes in the ORF or OE Comprehension measures are recommended at this time.

Multiple choice comprehension passages.

Bessie Coleman

Four items are redundant, having difficulty levels within .01 of one another. Items 1 and 10 are roughly equivalent, as are items 5 and 8. This redundancy reduces the information provided by the test to some degree. Again, we recommend either re-writing one of the items in each item pair in an attempt to differentiate the two better or replacing one of the items in each

item pair with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

The Stag

Items 14 and 15 have difficulty levels within .01 of one another. Again, we recommend either re-writing one of the items in each item pair in an attempt to differentiate the two better or replacing one of the items in each item pair with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

Vocabulary

IRT analysis indicates that four items of the Vocabulary measure are poorly fitted. Three of the items (Item #3, Mean Square Outfit = 1.65; Item #9, Mean Square Outfit = 1.88; and Item #10, Mean Square Outfit = 1.65) are more difficult than the other items on the test, while Item #11 (Mean Square Outfit = 1.88), is much less difficult than the rest of the items on the test. An analysis of distractors indicates that all four items are appropriately differentiating students by skill level. Thus, we recommend that the items be retained but that the district continues to monitor their functioning in future test administrations. Additionally, once again, the Vocabulary measure as a whole appears to be too easy for the students in this grade. More difficult items should be added to the measure, or the existing items should be revised to make them more challenging.

Grade 6 Measures

ORF and open-ended comprehension.

The sixth-grade ORF and OE Comprehension measures appear to be functioning well. They discriminate reliably between students of different skill levels. No changes in any of the them are recommended at this time.

Multiple choice comprehension passages.

Beowulf

IRT analysis indicates that Item 13 is too easy. In addition, it is fitted poorly (Mean Square Outfit = 0.32). Item 4 is too difficult and is poorly fitting (Mean Square Outfit = 3.90). In addition, five items are redundant, having difficulty levels within .01 of one another. Items11 and 14 are roughly equivalent, as are items 5, 9, and 15. Again, we recommend either re-writing one of the items in each item group in an attempt to differentiate them better or replacing one of the items in each item group with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

The Making of Shrek

Although IRT analysis indicates that Item 10 (Mean Square Outfit = 0.49) is just outside the generally acceptable range of .05 – 1.5, analysis of distractors indicates that it is appropriately differentiating students by skill level. Therefore, we recommend that the item be retained but that the district continue to monitor its functioning in future test administrations.

Vocabulary

#14 of the Vocabulary measure (Mean Square Outfit = 1.60). An analysis of distractors indicates that both items are appropriately differentiating students by skill level. Therefore, we

recommend that the items be retained but that the district continue to monitor their functioning in future test administrations.

Grade 7 Measures

All seventh-grade measures appear to be functioning well. Descriptive statistics indicate that they discriminate reliably between students of different skill levels and the Misfit statistics reflect no items functioning inappropriately. Distractor analysis indicates that all multiple choice items are appropriately differentiating between students of differing skill levels. No changes in any of the seventh-grade measures are recommended at this time.

Grade 8 Measures

All eighth-grade measures appear to be functioning well, although revising two items on each of the MC Comprehension measures might result in even more information being provided by the tests. Descriptive statistics indicate that all eighth-grade measures discriminate reliably between students of different skill levels and the Misfit statistics reflect no items functioning inappropriately. However, two items on each of the two MC Comprehension measures are redundant (Items 6 and 15 on the *P.T. Barnum* passage and Items 11 and 12 on the *Swiss Family Robinson* passage have the same difficulty level). Again, we recommend either re-writing one of the items in each item pair in an attempt to differentiate them better or replacing one of the items in each item pair with an item piloted in the original instrument development and continuing to evaluate the functioning of these items in the next cycle of testing.

Grade 9 Measures

This was the first year the district included tests at the ninth-grade level. All measures used for ninth-grade were appropriately differentiating students of different skill levels. However, because the measures were being piloted, they included more items than the measures

at the other grade levels. To bring the ninth-grade measures into line with the rest of the district assessments, the Vocabulary test needs to be abbreviated from 36 to 25 items, the MC Comprehension Test needs to be abbreviated from 25 to 15 items, and the Maze needs to be abbreviated from 60 to 30 items. Table 33 lists the items from the Vocabulary, MC Comprehension, and Maze measures that we recommend removing from the tests, based on IRT analyses.

Table 44
Items Recommended for Removal from Grade 9 Measures

Measure	Redundant with other items	Poorly fitting (Mean Square Outfit < 0.5 or > 1.5
Vocabulary	3, 8, 11, 14, 16, 18, 19, 20	6, 10, 22
MC Comprehension	3, 6, 7, 9, 13, 15, 16, 17, 21	24
Maze	4, 7, 8, 10, 11, 13, 14, 15, 19, 29, 30, 32, 33, 35, 38, 39, 44, 52, 53, 54, 55, 56, 57, 59	5, 16, 17, 18, 31, 40

Overall Recommendations

The ceiling effect occurring with the Vocabulary measure at grades 3, 4, 5, and 6 indicates that the measure currently is not providing much useful information about student skill across a very wide range. The measure would be improved by adding more challenging words or perhaps moving each of these grade level tests to the next grade down, thereby providing more of a 'stretch' for students who are taking these tests at the end of their academic year. Although such a change would cause a lowering of the district norm scores, this change might be worth considering.

District Decisions to Alter the Reading Assessment for the 2004- 2005 Assessment Cycle

The District Assessment Committee decided to follow all recommendations for re-writing or replacing poorly fitted multiple choice items listed above. In addition, the committee decided

to re-write the Grade 8 ORF passages by increasing the readability of the first 100 words in each passage in an attempt to bring the difficulty of the Grade 8 ORF passages more in line with the rest of the passages. To address concerns about the ceiling effect occurring with the Vocabulary measures at grades 3, 4, 5, and 6, the committee decided to increase the difficulty of each of those measures by replacing the easiest words on each grade level test with more difficult items from the vocabulary item bank piloted in previous years. To address two separate concerns (the need to reduce the cost of administering the District Reading Assessment and the ongoing problem with the Open Ended Comprehension questions not providing significant additional information that was not already supplied by other measures), the committee decided to drop the Open Ended Comprehension questions from every grade level. Table 45 presents the assessments that will be given during the 2004 – 2005 assessment cycle.

Table 45
District Reading Assessment Sub-Tests Planned for 2004 - 2005

Sub-test	Kindergarten			1st Grade			2nd Grade			3rd–8th Grade			9th Grade		
	F	W	S	F	W	S	F	W	S	F	W	S	F	W	S
Letter Naming		4	4	4											
Letter Sounds		4	4	4	4	4									
Phoneme Segmenting		4	4	4	4	4									
High Frequency Words					4	4	4		4						
2nd Grade Word Reading							4		4						
2nd Grade Sentence Reading							4	4							
ORF					4	4	4		4	4		4			
MC Comprehension					4	4	4		4	4		4	4		4
Vocabulary										4		4	4		4
Maze Comprehension													4		4

The current district reading assessments described in this technical report represent a rich source of information for students, parents, teachers, and administrators. Twice yearly assessments with individual scores as well as local norm-group information shared openly with all relevant stakeholders can allow students, parents, teachers, and administrators to make data-based decisions and monitor progress towards achieving skill attainment goals. In addition, longitudinal tracking of student progress provides one useful source of information to help the district evaluate the effectiveness of instructional materials and programs. As the district continues to develop and refine their reading assessments, extending them more fully into the high school years, the wealth of information available to teachers will continue to grow.

References

Alonzo, J., & Tindal, G. (2004a). *Analysis of reading fluency and comprehension measures for fourth-grade students (Tech. Rep. No. 27).* Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Alonzo, J., & Tindal, G. (2004b). *Analysis of reading fluency and comprehension measures for first-grade students (Tech. Rep. No. 25).* Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Alonzo, J., & Tindal, G. (2004c). *Analysis of reading fluency and comprehension measures for sixth-grade students (Tech. Rep. No. 24).* Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Alonzo, J., & Tindal, G. (2004d). *Analysis of reading fluency and comprehension measures for seventh-grade students (Tech. Rep. No. 23).* Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Fry, F., Kress, J., & Fountonkidis, D. (2000). *The Reading Teacher's Book of Lists*

Ketterlin-Geller, L.R., & Tindal, G. (2004). *Analysis of reading fluency and comprehension measures for 3rd-grade students (Tech. Rep. No. 22).* Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Linacre, J. M. (3.48). WINSTEPS (Version 3.48) [Computer Software]. Chicago, IL: Author.

Shoemaker, B. J., & Alonzo, J. (n.d.). Every 4j child a reader. Manuscript submitted for publication.