Colorado
English Language Acquisition Assessment Program

2010 Technical Report

Submitted to the
Colorado Department of Education

July 31, 2010

CTB McGraw-Hill

Table of Contents

Appendices

APPENDIX A: Item Analysis Results

APPENDIX B: Comparison of CELApro 2008 and CELApro 2009/2010 Anchor Parameters

APPENDIX C: TCC & SEM Plots by Grade Span

APPENDIX D: Equating Results for Grade Spans

APPENDIX E: Raw Score to Scale Score Tables

APPENDIX F: List of Colorado Standards Directly Assessed by the CELApro

List of Figures

List of Tables

Overview

The first administration of the Colorado English Language Acquisition Assessments (CELApro) occurred in spring 2006. At that time, the assessments were identical to CTB's *LAS Links* ® Form A, except for customized Colorado test book covers and answer sheets.

LAS Links (Form A) continues to provide a solid foundation for all the CELApro tests. The *LAS Links* assessments were developed from a framework that reflects sound principles of second-language acquisition (Schmidt, 2001; Savignon, 1997, 1972; Bachman & Palmer, 1996; O'Malley & Valdez Pierce, 1996; Chamot & O'Malley, 1994; Bachman, 1990). Each *LAS Links* test consists of four separately scored sections (Reading, Writing, Listening, and Speaking). In addition to these four component scores, all of the Listening and Speaking items are combined to produce an Oral score, and selected Reading and Listening items are combined to yield a Comprehension score.

Approximately 30,000 students participated in the field test, item analysis, and calibration of *LAS Links* Form A, which was calibrated and scaled using item-response theory and a common-item equating design to place all grade levels on a common scale and to ensure that skill area scores have the same meaning across forms, grades, and years.

The *LAS Links* tests are aligned to CTB/McGraw-Hill's English Language Proficiency Assessment Standards (ELPAS), which were developed to include key standards from the national ESL and TESOL standards and from several state ESL standards. In order to match the Colorado standards, the five *LAS Links* grade spans were modified and reduced to four grade spans for the 2007 CELApro tests. In 2008, the K–2 grade span was divided into three separate grade spans, resulting in the current six CELApro grade spans: K, 1, 2, 3–5, 6–8, and 9–12. Additional items were added in 2008 to the Reading section in Grades K, 1, and 2 and to the Listening and Writing sections in all six grade spans. Table 1 shows a comparison of grade spans by year.

Table 1. Comparison of LAS Links and CELApro Grade Spans

Grade Spans		
LAS Links	CELApro 2007	CELApro 2008–2010
K–1	K–2	K
		1
2–3		2
4–5	3–5	3–5
6–8	6–8	6–8
9–12	9–12	9–12

The 2010 CELApro tests are identical to those administered in 2008 and 2009. K, 1, 2, and 3–5 have a scannable book; the other grade spans have a reusable test book and a scannable answer book. The Speaking items and the Writing constructed-response (CR) items appear only in the answer sheet for the upper two grade spans.

Part 1: Standards

The Colorado English Language Acquisition Assessment (CELApro) is the language proficiency assessment used for classifying and monitoring the progress of Colorado English Language Learners (ELLs) in the acquisition of English. *LAS Links* Form A assessments form the core of the CELApro tests.

The CELApro assessment measures the competencies necessary for successful social and academic language use in four major modalities—Listening, Speaking, Reading, and Writing—along a continuum of five proficiency levels: Beginning, Early Intermediate, Intermediate, Proficient, and Advanced. The assessment takes into account the students' maturation and cognitive skills by providing age appropriate tests covering six grade spans: K, 1, 2, 3–5, 6–8, and 9–12.

A combination of item types—constructed-response (CR) and multiple choice (MC) items—provide a variety of ways for students to demonstrate proficiency and to maintain reasonable testing times. Constructed-response items assess the productive domains of Speaking and Writing, whereas the multiple-choice items assess the receptive domains of Listening, Reading, and Writing (grammar). The variety of item types ensures measurement of the full spectrum of possible tasks required for each language subskill and allows for the interpretation of the results in multiple ways.

Alignment Studies

An important indicator of the validity of a standardized test is the degree of *alignment* (i.e., the match) between the state English language development (ELD) standards and the test content. In developing standardized tests, test items are written to cover as many standards as possible.

Colorado has four general standards for English language learners, organized by modality (Listening, Speaking, Reading, and Writing) and applicable at all grade levels. The standards specify general skills in social and academic language:

- **Standard 1:** English Language Learners listen for information and understanding, using a variety of sources, for academic and social purposes.

- **Standard 2:** English Language Learners speak to convey information and understanding, using a variety of sources, for academic and social purposes.

- **Standard 3:** English Language Learners read for information and understanding, using a variety of sources, for academic and social purposes.

- **Standard 4:** English Language Learners write to convey information and understanding, using a variety of sources, for academic and social purposes.

A detailed description of the standards, by grade and proficiency level, is provided in Appendix F.

In order to increase the alignment of CELApro to the Colorado ELD standards, additional test items were written for the 2008 tests to assess individual standards that were not already assessed by *LAS Links* items.

CTB conducted an alignment analysis of the 2008 CELApro assessments to evaluate the match between the test and the standards. In performing an alignment, it is sometimes necessary to eliminate some standards because they cannot be easily assessed by a standardized test. For example, a standard may require an extended process outside of the test situation, as in the steps for writing a research paper, or it may specify instructional strategies rather than student skills, or it may specify parameters outside of the testing situation, such as "participate in group discussions." Of the 397 Colorado ELD standards, 104 were eliminated as nonassessable. In performing the CELApro alignment, the raters independently matched items to all of the assessable standards on the basis of direct, indirect, or partial alignment. The test item numbers were then entered into the cells of the matching standards. The degree of alignment was calculated by adding up the number of assessable standards that were measured by at least one CELApro item. All of the standards are assessed by at least one test item.

This alignment was reviewed by CTB and the Colorado Department of Education (CDE) at a meeting in April 2007. CTB then conducted a final review with a committee of English Language Acquisition experts finalizing the alignment shown in Table 2. This table also reflects the current alignment, because the CELApro tests were unchanged from 2008 to 2010.

Table 2. Item Alignment Percentages by Grade Span

	K–2		3–5		6–8		9–12	
Listening	14/14	100	15/15	100	14/14	100	15/15	100
Beginning	4/4		5/5		5/5		5/5	
Intermediate	5/5		5/5		5/5		5/5	
Advanced	5/5		5/5		4/4		5/5	
Speaking	14/14	100	13/13	100	12/12	100	11/11	100
Beginning	4/4		4/4		5/5		4/4	
Intermediate	5/5		4/4		4/4		4/4	
Advanced	5/5		5/5		3/3		3/3	
Reading	16/16	100	15/15	100	14/14	100	12/12	100
Beginning	5/5		4/4		3/3		3/3	
Intermediate	7/7		6/6		5/5		5/5	
Advanced	4/4		5/5		6/6		4/4	

Table 2. Item Alignment Percentages by Grade Span (continued)

	K–2		3–5		6–8		9–12	
Writing	13/13	100	14/14	100	9/9	100	7/7	100
Beginning	2/2		4/4		3/3		2/2	
Intermediate	5/5		5/5		3/3		3/3	
Advanced	6/6		5/5		3/3		2/2	

Part 2: Test Development

The 2010 CELApro tests are identical to the 2008 and 2009 tests and consist of both *LAS Links* items and items owned by the Colorado Department of Education. For the two upper grade spans (Grades 6–8 and 9–12), the organization of the CELApro tests are identical to the corresponding *LAS Links* assessments. The reconfigured tests for the two lower grade spans (Grades K–2 and 3–5) were created using selected items from the *LAS Links* assessments for the appropriate grades. The lowest grade span was also broken out into separate tests for Kindergarten, Grade 1, and Grade 2. All K–2 students take the same Listening and Speaking items but some different Reading and Writing items. All of these items were written by writers with experience or training in the areas being tested. Before writing items, all writers went through extensive training and were instructed to:

- Study each standard to be assessed.
- Decide what is important for the student to know and do to demonstrate mastery of the standard. Avoid the trivial.
- Write the item so that it focuses on the particular content or skill to be assessed.
- Develop answer choices that relate logically to the stem and standard. The correct response should be clear to students who have mastered the concept or skill. The distractors should be clearly wrong to students who have mastered the content or skill. Test items should not be "tricky" or contain information unfamiliar to most students.
- Provide documentation from source material (e.g., photocopies of encyclopedia entries and other reliable reference materials) to verify that all information included in the stimulus and item is correct. All factual statements in stimuli, stems, and correct responses must be checked against reliable sources. Distractors also should be verified as incorrect.
- Use appropriate subject matter. Refrain from explicit references to or descriptions of alcohol or drug abuse, sex, or vulgar language. Exercise caution when developing religious, political, social, or philosophical issues as subject matter. Individual beliefs should not influence content.
- Avoid using very controversial material. Large-scale (national, state, or district) assessments are administered to student populations with different experiences and beliefs.
- Verify that the item is free of content that could be offensive, insensitive, stereotypical, or that introduces other types of bias.
- Check that the content of the stimulus and/or the item is developmentally and age appropriate for the students being tested.
- Write a range of items representing all levels of proficiency in English within a specific standard.

The tests have been structured to comprehensively assess the four language skills of Speaking, Listening, Reading, and Writing. Comprehension is assessed using selected Listening and Reading items. A combination of constructed-response, dichotomous constructed-response (correct or incorrect), and multiple-choice items is used to provide diverse opportunities for students to demonstrate proficiency and to maintain reasonable testing times. Constructed-response items are used to assess the productive domains of Speaking and Writing, whereas the multiple-choice items are used to assess the receptive domains of Listening, Reading, and the Writing Use Conventions subtest. The structure of the 2010 CELApro is shown in Table 3.

Table 3. 2010 CELApro Test Structure

Content	Grade Span	Sub-Content	Item Type	Items	Score Points	CR/DCR Items Scored By	Administration
Speaking 20 items, 41 pts	4 grade spans: K–2, 3–5, 6–8, 9–12	Speak Words	DCR*	10	10	Local Test Administrator	Individual
		Sentences	CR	5	15		
		Conversation	CR	4	12		
		Tell a Story	CR	1	4		
Listening K=21 items, 20 pts** 1–2=21 items, 21 pts 3–12=23 items, 23 pts	K	Listen for Information	MC	11	11	Not Applicable	Individual/Group
		Listen in the Classroom	MC	5	5		
		Listen & Comprehend	MC	4	4		
	1–2	Listen for Information	MC	11	11		
		Listen in the Classroom	MC	6	6		
		Listen & Comprehend	MC	4	4		
	3–12	Listen for Information	MC	10	10		
		Listen in the Classroom	MC	9	9		
		Listen & Comprehend	MC	4	4		
Reading K=31 items, 31 pts 1–2=36 items, 36 pts 3–12=35 items, 35 pts	K	Analyze Words	MC	11	11	Not Applicable	Individual
		Read Words	MC	10	10		
		Understanding	MC	10	10		
	1–2	Analyze Words	MC	11	11		
		Read Words	MC	10	10		
		Understanding	MC	15	15		
	3–12	Analyze Words	MC	10	10		Group
		Read Words	MC	10	10		
		Understanding	MC	15	15		
Writing K–1=25 items, 35 pts 3–5=26 items, 37 pts 2, 6–12=25 items, 36 pts	K–1	Conventions	MC	20	20	CTB Handscoring	Group (or Individual for K)
		Write About	CR	2	6		
		Write Why	CR	3	9		
	3–5	Conventions	MC	21	21	CTB Handscoring	Group
		Write About	CR	2	6		
		Write Why	CR	2	6		
		Write in Detail	CR	1	4		
	2, 6–12	Conventions	MC	20	20	CTB Handscoring	Group
		Write About	CR	2	6		
		Write Why	CR	2	6		
		Write in Detail	CR	1	4		
Oral K–2=40 items, 61 pts 1–2=41 items, 62 pts 3–12=43 items, 64 pts	K–2	Listening and Speaking	MC	20	20	Local Test Administrator	Not Applicable
			SCR	10	31		
			ECR	10	10		
	1–2	Listening and Speaking	MC	21	21		
			SCR	10	31		
			ECR	10	10		
	3–12	Listening and Speaking	MC	23	23		
			SCR	9	27		
			ECR	1	4		
			CR	10	10		
Comprehension K=39 items, 39 pts 1–2=45 items, 45 pts 3–5=48 items, 48 pts 6–12=50 items, 50 pts	K	Listening and Reading	MC	39	39	Not Applicable	Not Applicable
	1–2	Listening and Reading	MC	45	45		
	3–5	Listening and Reading	MC	48	48		
	6–12	Listening and Reading	MC	50	50		

*KEY: DCR=Dichotomous CR; SCR=Short CR; ECR=Extended CR

** There were 21 items in the Kindergarten Listening test, but one item was suppressed prior to scoring. This item was found to be performing poorly.

Item Review and Test Fairness

All items are expected to be fair for all examinees. Various procedures are employed to review item bias. Once the items are developed, they must go through a series of content and bias reviews and analyses prior to being selected as part of the item pool. A content and bias review has two purposes: to ensure that the items are grade level appropriate and to ensure that any sensitivity issues are identified and addressed. Grade level appropriateness is evaluated by grade level teachers who possess the on-the-ground knowledge of how content is taught in the classroom. Sensitivity reviews ensure that items are free of offensive, disturbing, or inappropriate language or content.

Content reviews and sensitivity and bias reviews were conducted on all operational items. The item review committees reviewed all operational items before the operational test administration.

Item Selection

In selecting items for the reconfigured CELApro tests in Grades K–2 and 3–5, the primary criterion was to meet the content specifications represented by test blueprints, while at the same time maintaining the desired statistical properties of *LAS Links*. This involved an iterative process in which test characteristic curves and standard errors were examined after each preliminary item selection. Selections were revised as necessary in order to obtain an acceptable match to the statistical properties of the previous *LAS Links* assessments at each grade level.

Minimizing Test Bias

The position of CTB/McGraw-Hill concerning test bias is based on two general propositions. First, students may differ in their background knowledge, cognitive and academic skills, language, attitudes, and values. To the degree that these differences are large, no one curriculum and no one set of instructional materials will be equally suitable for all. Therefore, no one test will be equally appropriate for all. Furthermore, it is difficult to specify what amount of difference can be called large and to determine how these differences will affect the outcome of a particular test.

Second, schools have been assigned the tasks of developing certain basic cognitive skills and supporting English language proficiency among all students. Therefore, there is a need for ELP tests that measure the common skills and bodies of knowledge that are common to English learners. The test publisher's task is to develop assessments that measure English language proficiency without introducing extraneous or construct-irrelevant elements in the performances on which the measurement is based. If these tests require that students have cultural specific knowledge and skills not taught in school, differences in performance among students can occur because of differences in student background and out-of-school learning. Such tests are measuring different things for different groups and can be called biased (Camilli & Shepard, 1994; Green, 1975). In order to lessen this bias, CTB/McGraw-Hill strives to minimize the role of the extraneous elements, thereby increasing the number of students for whom the test is appropriate. Careful attention is given during the test construction process to lessen the influence of these elements for large numbers of students. Unfortunately, in some cases these elements may continue to play a substantial role.

Four measures were taken to minimize bias in the *LAS Links* assessments. The first was based on the premise that careful editorial attention to validity is an essential step in keeping bias to a minimum. Bias can occur only if the test is measuring different things for different groups. If the test entails irrelevant skills or knowledge, however common, the possibility of bias is increased. Thus, careful attention was given to content validity during the item-writing and item-selection process.

The second way bias was minimized was by following the McGraw-Hill guidelines designed to reduce or eliminate bias. Item writers were directed to the following published guidelines: *Guidelines for Bias-Free Publishing* (MacMillan/McGraw-Hill, 1993a) and *Reflecting Diversity: Multicultural Guidelines for Educational Publishing Professionals* (Macmillan/McGraw-Hill, 1993b). Developers reviewed *LAS Links* Assessment materials with these considerations in mind. Such internal editorial reviews were conducted by at least four separate people: a content editor, who directly supervised the item writers; the project director; a style editor; and a proofreader. The final test built from the tryout materials was again reviewed by at least these same people.

In the third effort to minimize bias, educational community professionals who represent various ethnic groups reviewed all *LAS Links* tryout materials. They were asked to consider and comment on the appropriateness of language, subject matter, and representation of groups of people.

It is believed that these three procedures both improve the quality of an assessment and reduce item and test bias. However, current evidence suggests that expertise in this area is no substitute for data. Reviewers are often wrong about which items perform differently between specific subgroups of students, apparently because some of their ideas about how students will react to items may be inaccurate (Camilli & Shepard, 1994; Sandoval & Mille, 1979; Scheuneman, 1984). Thus, a fourth method for minimizing bias, an empirical approach, was also used to identify potential sources of item bias. For language tests, these are differential item functioning (DIF) studies, since criterion-related validities are essentially unobtainable for such tests. DIF studies include a systematic item analysis to determine whether examinees with the same underlying level of ability have the same probability of getting the item correct. Items identified with DIF are then examined to determine whether item performance differences between identifiable subgroups of the population are due to extraneous or construct-irrelevant information making the items unfairly difficult. The inclusion of these items is minimized in the test development process. DIF studies have been routinely done for all major test batteries published by CTB/McGraw-Hill after 1970. Differential item functioning of the *LAS Links* assessment tryout items was assessed for students identified as males and females at each grade level in which the items were administered. In most cases, each item was administered at two grade spans.

Because *LAS Links* was built using item response theory, DIF analyses that capitalized on the information and item statistics provided by this theory were implemented. There are several IRT-based DIF procedures, including those that assess the equality of item parameters across groups (Lord, 1980), and those that assess area differences between item characteristic curves (Linn, Levine, Hastings, & Wardrop, 1981; Camilli & Shepard, 1994). However, these procedures require a minimum of 800 to 1000 cases in each group of comparison to produce reliable and consistent results. In contrast, the Linn-Harnisch procedure (Linn & Harnisch, 1981) utilizes the information provided by the three-parameter IRT model but requires fewer cases. This was the procedure used to complete the gender DIF studies for the *LAS Links* field test data.

Part 3: Tested Population

A total of 95,547 students participated in the 2010 CELApro testing. Students in Kindergarten and Grade 1 formed the largest groups of examinees (12,738 and 12,821 respectively), with numbers generally decreasing at successive grade levels. The number of male examinees was slightly greater than the number of females at each grade level. The examinee counts by grade and gender are shown in Table 4. Note that not all students completed all four of the CELApro content areas, so these numbers differ from those that appear in some of the subsequent tables within this report.

Table 4. Examinee Counts by Grade and Gender

Grade	Number of Examinees			Total
	Females	Males	Not Specified	
Kindergarten	6138	6597	3	12738
1	6252	6567	2	12821
2	5831	6257	1	12089
3	5401	5730	2	11133
4	4695	4949	2	9646
5	3535	4029	0	7564
6	2995	3339	1	6335
7	2440	2793	2	5235
8	2043	2483	4	4530
9	1869	2448	1	4318
10	1575	1851	0	3426
11	1370	1635	3	3008
12	1290	1413	1	2704
Total	45434	50091	22	95547

Student ethnicity and home language is summarized by grade span in Tables 5 and 6.

Table 5. Ethnicity by Grade Span

Ethnicity	Grade Span							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
American Indian/ Alaska Native	131	0.14	122	0.13	106	0.11	58	0.06
Asian/Pacific Islander	2856	2.99	1895	1.98	989	1.04	1151	1.2
Black	979	1.02	651	0.68	378	0.40	583	0.61
Hispanic	31001	32.45	24307	25.44	13894	14.54	10952	11.46
White	2678	2.80	1365	1.43	727	0.76	710	0.74
Not Specified	3	0.00	3	0.00	6	0.01	2	0

Table 6. Home Language (193 Languages Represented)

Language	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Afar	1	0.00	0	0.00	0	0.00	2	0.00
Afrikaans	8	0.01	7	0.01	1	0.00	3	0.00
Akan	18	0.02	6	0.01	2	0.00	13	0.01
Albanian	11	0.01	10	0.01	3	0.00	1	0.00
Amharic	222	0.23	125	0.13	53	0.06	76	0.08
Anuak	1	0.00	3	0.00	0	0.00	1	0.00
Apache	0	0.00	1	0.00	0	0.00	0	0.00
Arabic	478	0.50	270	0.28	143	0.15	139	0.15
Arapaho	2	0.00	1	0.00	1	0.00	0	0.00
Armenian	16	0.02	7	0.01	3	0.00	4	0.00
Assamese	8	0.01	9	0.01	4	0.00	0	0.00
Assyrian	1	0.00	0	0.00	0	0.00	0	0.00
Awadhi	0	0.00	1	0.00	0	0.00	0	0.00
Azerbaijani	1	0.00	0	0.00	0	0.00	0	0.00
Bambara	3	0.00	5	0.01	0	0.00	1	0.00
Bashkir	1	0.00	1	0.00	2	0.00	1	0.00
Bassa	0	0.00	2	0.00	0	0.00	0	0.00
Bemba	2	0.00	0	0.00	0	0.00	0	0.00
Bengali	23	0.02	9	0.01	2	0.00	4	0.00
Berber	2	0.00	0	0.00	0	0.00	0	0.00
Bihari	1	0.00	1	0.00	0	0.00	0	0.00
Bosnian	58	0.06	30	0.03	11	0.01	8	0.01
Bulgarian	17	0.02	6	0.01	6	0.01	6	0.01
Burmese	66	0.07	67	0.07	51	0.05	65	0.07
Cahuilla	0	0.00	0	0.00	1	0.00	0	0.00
Cebuano	1	0.00	6	0.01	3	0.00	1	0.00
Chamorro	5	0.01	3	0.00	5	0.01	3	0.00
Chechen	1	0.00	0	0.00	0	0.00	0	0.00
Chinese, Cantonese	151	0.16	86	0.09	25	0.03	42	0.04
Chinese, Hakka	1	0.00	7	0.01	5	0.01	11	0.01
Chinese, Mandarin	341	0.36	150	0.16	88	0.09	132	0.14
Chinese, Min Nan	1	0.00	0	0.00	0	0.00	0	0.00
Chinese, Wu	4	0.00	0	0.00	0	0.00	1	0.00
Choctaw	0	0.00	0	0.00	1	0.00	0	0.00
Chuukese	11	0.01	9	0.01	5	0.01	7	0.01
Chuvash	0	0.00	6	0.01	0	0.00	0	0.00
Cora	26	0.03	18	0.02	28	0.03	7	0.01
Cree	0	0.00	1	0.00	0	0.00	0	0.00
Croatian	3	0.00	3	0.00	2	0.00	1	0.00

Table 6. Home Language (continued)

Language	Test Level							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Czech	19	0.02	10	0.01	6	0.01	4	0.00
Danish	14	0.01	11	0.01	3	0.00	2	0.00
Dari	4	0.00	2	0.00	3	0.00	3	0.00
Deccan	3	0.00	2	0.00	0	0.00	1	0.00
Dinka	11	0.01	6	0.01	5	0.01	1	0.00
Dutch	16	0.02	11	0.01	4	0.00	4	0.00
Edo	1	0.00	0	0.00	0	0.00	0	0.00
Efik	1	0.00	0	0.00	0	0.00	0	0.00
English	10	0.01	1	0.00	4	0.00	2	0.00
Eskimo	0	0.00	0	0.00	2	0.00	0	0.00
Estonian	1	0.00	1	0.00	0	0.00	2	0.00
Ewe	6	0.01	4	0.00	2	0.00	2	0.00
Faroese	1	0.00	2	0.00	0	0.00	1	0.00
Farsi, Eastern	73	0.08	37	0.04	10	0.01	16	0.02
Farsi, Western	31	0.03	23	0.02	6	0.01	11	0.01
Fijian	0	0.00	0	0.00	1	0.00	1	0.00
Finnish	12	0.01	2	0.00	1	0.00	0	0.00
Flemish	1	0.00	0	0.00	0	0.00	0	0.00
French	131	0.14	87	0.09	59	0.06	98	0.10
French Cree	1	0.00	2	0.00	0	0.00	0	0.00
French Creole	12	0.01	17	0.02	10	0.01	6	0.01
Fulfulde, Nigerian	5	0.01	5	0.01	4	0.00	2	0.00
Ga	0	0.00	3	0.00	0	0.00	1	0.00
Ganda	3	0.00	2	0.00	2	0.00	3	0.00
Georgian	1	0.00	0	0.00	0	0.00	0	0.00
German	112	0.12	51	0.05	31	0.03	24	0.03
Grebo	0	0.00	1	0.00	0	0.00	1	0.00
Greek	7	0.01	3	0.00	1	0.00	1	0.00
Gujarati	17	0.02	4	0.00	4	0.00	0	0.00
Haitian, Creole French	2	0.00	4	0.00	1	0.00	0	0.00
Hausa	0	0.00	1	0.00	2	0.00	0	0.00
Hawaiian	1	0.00	0	0.00	1	0.00	1	0.00
Hazaragi	0	0.00	0	0.00	0	0.00	1	0.00
Hebrew	15	0.02	7	0.01	7	0.01	7	0.01
Hindi	81	0.08	42	0.04	17	0.02	6	0.01
Hmong	185	0.19	132	0.14	82	0.09	74	0.08
Hungarian	11	0.01	5	0.01	2	0.00	0	0.00
Icelandic	1	0.00	0	0.00	0	0.00	0	0.00
Igbo	11	0.01	7	0.01	5	0.01	5	0.01

Table 6. Home Language (continued)

Language	Test Level							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Ilocano	1	0.00	2	0.00	2	0.00	2	0.00
Indonesian	50	0.05	29	0.03	18	0.02	16	0.02
Italian	17	0.02	7	0.01	5	0.01	11	0.01
Japanese	78	0.08	56	0.06	11	0.01	20	0.02
Kanjobal	17	0.02	18	0.02	9	0.01	3	0.00
Kannada	12	0.01	1	0.00	0	0.00	0	0.00
Karelian	0	0.00	1	0.00	0	0.00	1	0.00
Karen	62	0.06	60	0.06	45	0.05	95	0.10
Kazakh	1	0.00	1	0.00	1	0.00	0	0.00
Keres, Eastern	0	0.00	1	0.00	1	0.00	0	0.00
Khmer	70	0.07	66	0.07	34	0.04	30	0.03
Kikuyu	0	0.00	1	0.00	0	0.00	0	0.00
Kinyarwanda	6	0.01	2	0.00	3	0.00	5	0.01
Koli, Kachi	1	0.00	0	0.00	0	0.00	0	0.00
Korean	305	0.32	210	0.22	114	0.12	104	0.11
Kosraen	0	0.00	0	0.00	0	0.00	1	0.00
Kpelle	0	0.00	1	0.00	0	0.00	0	0.00
Krahn	2	0.00	3	0.00	0	0.00	7	0.01
Krio	3	0.00	2	0.00	0	0.00	7	0.01
Kru Languages	0	0.00	1	0.00	2	0.00	2	0.00
Kru Pidgin English	0	0.00	0	0.00	0	0.00	3	0.00
Kurdi/Kurdish Bandinani	21	0.02	15	0.02	7	0.01	0	0.00
Lakota	3	0.00	2	0.00	1	0.00	0	0.00
Lao	60	0.06	57	0.06	23	0.02	19	0.02
Latvian	2	0.00	1	0.00	1	0.00	1	0.00
Lezgi	0	0.00	0	0.00	1	0.00	0	0.00
Liberian English	7	0.01	5	0.01	1	0.00	4	0.00
Lingala	3	0.00	1	0.00	3	0.00	1	0.00
Lithuanian	8	0.01	6	0.01	5	0.01	4	0.00
Lwo	0	0.00	0	0.00	1	0.00	0	0.00
Maay	11	0.01	9	0.01	4	0.00	4	0.00
Macedonian	1	0.00	0	0.00	0	0.00	0	0.00
Malagasy	1	0.00	0	0.00	0	0.00	0	0.00
Malay	10	0.01	2	0.00	4	0.00	0	0.00
Malayalam	22	0.02	10	0.01	4	0.00	1	0.00
Mandinka	14	0.01	8	0.01	2	0.00	5	0.01
Maniinkakan, Western	0	0.00	0	0.00	1	0.00	0	0.00

Table 6. Home Language (continued)

Language	Test Level							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Marathi	18	0.02	4	0.00	1	0.00	0	0.00
Marshallese	23	0.02	13	0.01	9	0.01	7	0.01
Maya	2	0.00	1	0.00	0	0.00	2	0.00
Mende	0	0.00	0	0.00	0	0.00	1	0.00
Mixteco, San Juan Mixtepec	0	0.00	0	0.00	0	0.00	1	0.00
Mongolian	40	0.04	30	0.03	14	0.01	25	0.03
Mono	1	0.00	0	0.00	0	0.00	0	0.00
Navajo	60	0.06	53	0.06	38	0.04	23	0.02
Nepali	108	0.11	104	0.11	93	0.10	141	0.15
Newari	1	0.00	0	0.00	1	0.00	0	0.00
Norwegian	3	0.00	1	0.00	0	0.00	1	0.00
Nuer	7	0.01	3	0.00	1	0.00	2	0.00
Nyanja	0	0.00	1	0.00	1	0.00	0	0.00
Oriya	7	0.01	0	0.00	0	0.00	0	0.00
Oromo, West-Central	31	0.03	20	0.02	8	0.01	28	0.03
Palauan	4	0.00	1	0.00	2	0.00	2	0.00
Panjabi, Eastern	18	0.02	13	0.01	12	0.01	9	0.01
Panjabi, Western	3	0.00	3	0.00	0	0.00	0	0.00
Pashto, Central	3	0.00	4	0.00	1	0.00	3	0.00
Pashto, Northern	5	0.01	1	0.00	2	0.00	2	0.00
Pashto, Southern	5	0.01	5	0.01	2	0.00	0	0.00
Pohnpeian	7	0.01	1	0.00	1	0.00	2	0.00
Polish	73	0.08	49	0.05	10	0.01	11	0.01
Pomo, Northern	1	0.00	0	0.00	0	0.00	0	0.00
Portuguese	36	0.04	26	0.03	11	0.01	16	0.02
Pulaar	0	0.00	2	0.00	2	0.00	2	0.00
Quiche, Central	0	0.00	0	0.00	1	0.00	2	0.00
Romani	0	0.00	0	0.00	0	0.00	1	0.00
Romanian	24	0.03	8	0.01	5	0.01	7	0.01
Rundi	21	0.02	18	0.02	8	0.01	14	0.01
Russian	364	0.38	259	0.27	140	0.15	145	0.15
Rwanda	2	0.00	4	0.00	4	0.00	1	0.00
Samoan	13	0.01	8	0.01	4	0.00	6	0.01
Sango	0	0.00	0	0.00	1	0.00	0	0.00
Serbian	9	0.01	5	0.01	1	0.00	3	0.00
Serbo-Croatian	10	0.01	6	0.01	11	0.01	9	0.01
Sesotho	4	0.00	2	0.00	0	0.00	0	0.00
Shona	2	0.00	1	0.00	1	0.00	0	0.00

Table 6. Home Language (continued)

Language	Test Level							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Sindhi	2	0.00	0	0.00	0	0.00	0	0.00
Sinhala	1	0.00	0	0.00	0	0.00	1	0.00
Sioux	0	0.00	0	0.00	1	0.00	0	0.00
Slovak	4	0.00	1	0.00	2	0.00	1	0.00
Slovenian	0	0.00	0	0.00	1	0.00	0	0.00
Somali	164	0.17	122	0.13	93	0.10	141	0.15
Soninke	0	0.00	1	0.00	0	0.00	0	0.00
Spanish	32127	33.62	24792	25.95	14169	14.83	11179	11.70
Spokane	1	0.00	0	0.00	0	0.00	0	0.00
Sundanese	0	0.00	0	0.00	1	0.00	2	0.00
Susu	1	0.00	0	0.00	2	0.00	1	0.00
Swahili	27	0.03	25	0.03	15	0.02	35	0.04
Swedish	17	0.02	11	0.01	2	0.00	1	0.00
Tagalog	106	0.11	72	0.08	54	0.06	43	0.05
Tamil	45	0.05	14	0.01	3	0.00	1	0.00
Telugu	67	0.07	25	0.03	5	0.01	1	0.00
Thai	30	0.03	25	0.03	14	0.01	22	0.02
Tibetan	4	0.00	2	0.00	1	0.00	5	0.01
Tigrigna	61	0.06	32	0.03	18	0.02	35	0.04
Tlingit	0	0.00	0	0.00	0	0.00	1	0.00
Tok Pisin	0	0.00	1	0.00	0	0.00	0	0.00
Tonga	1	0.00	2	0.00	3	0.00	2	0.00
Tongan	0	0.00	2	0.00	0	0.00	0	0.00
Tonkawa	0	0.00	0	0.00	0	0.00	1	0.00
Tsonga	1	0.00	0	0.00	0	0.00	2	0.00
Tswana	0	0.00	0	0.00	0	0.00	1	0.00
Turkish	32	0.03	31	0.03	11	0.01	13	0.01
Turkmen	0	0.00	1	0.00	0	0.00	0	0.00
Twi	52	0.05	20	0.02	14	0.01	26	0.03
Ukrainian	35	0.04	34	0.04	25	0.03	23	0.02
Urdu	56	0.06	37	0.04	13	0.01	13	0.01
Ute	7	0.01	13	0.01	15	0.02	19	0.02
Uzbek	7	0.01	2	0.00	2	0.00	1	0.00
Vengo	0	0.00	0	0.00	0	0.00	1	0.00
Vietnamese	790	0.83	495	0.52	203	0.21	225	0.24
Welsh	0	0.00	1	0.00	0	0.00	0	0.00
Wolof	3	0.00	1	0.00	1	0.00	1	0.00
Yapese	2	0.00	1	0.00	0	0.00	0	0.00

Table 6. Home Language (continued)

Language	Test Level							
	Grades K–2		Grades 3–5		Grades 6–8		Grades 9–12	
	N	%	N	%	N	%	N	%
Yoruba	8	0.01	8	0.01	6	0.01	4	0.00
Zuni	1	0.00	0	0.00	0	0.00	0	0.00
Unspecified/No language code match	18	0.01	13	0.01	18	0.01	34	0.03
Total	37648	39.29	28343	29.62	16100	16.77	13456	14.01

Because some students required accommodations in order to access the items, the following accommodations were available:

- Braille
- Large Print
- Use of a Scribe to Record Responses
- Signing
- Use of Assistive Communicative Devices
- Use of Approved Nonstandard Accommodations
- Oral Presentation

These accommodations are summarized, by content area and grade, in Tables 7 to 10.

Table 7. Speaking Accommodations by Grade

Grade	Speaking Accommodations Provided							Total
	None	Braille	Large Print	Signing	Assistive Communication Device	Approved Nonstandard Accommodation	Not Specified	
K	12716	0	2	1	1	0	18	12738
1	12792	1	1	2	0	1	24	12821
2	12061	0	4	3	2	0	19	12089
3	11108	0	6	4	1	0	14	11133
4	9624	1	2	1	3	1	14	9646
5	7542	1	1	2	1	2	15	7564
6	6315	1	1	7	1	0	10	6335
7	5222	1	0	3	1	1	7	5235
8	4519	0	1	2	0	0	8	4530
9	4314	0	0	2	0	0	2	4318
10	3417	1	0	1	0	0	7	3426
11	2998	0	0	0	1	0	9	3008
12	2688	1	0	5	0	0	10	2704
TOTAL	95316	7	18	33	11	5	157	95547

Table 8. Listening Accommodations by Grade

Grade	Listening Accommodations Provided							Total
	None	Braille	Large Print	Signing	Assistive Communication Device	Approved Nonstandard Accommodation	Not Specified	
K	12714	0	2	1	0	1	20	12738
1	12790	2	1	2	0	2	24	12821
2	12060	0	4	3	1	0	21	12089
3	11104	0	6	3	1	5	14	11133
4	9624	1	2	1	1	3	14	9646
5	7541	1	2	2	1	2	15	7564
6	6315	1	2	3	2	1	11	6335
7	5224	1	0	2	0	1	7	5235
8	4519	0	1	1	0	1	8	4530
9	4313	0	0	2	0	0	3	4318
10	3416	1	0	1	0	1	7	3426
11	2995	0	0	0	1	0	12	3008
12	2688	1	0	4	0	1	10	2704
TOTAL	95303	8	20	25	7	18	166	95547

Table 9. Reading Accommodations by Grade

Grade	Reading Accommodation Provided								Total
	None	Braille	Large Print	Scribe	Signing	Assistive Communication Device	Approved Nonstandard Accommodation	Not Specified	
K	12709	0	2	4	1	0		22	12738
1	12785	2	1	2	5	0	1	25	12821
2	12057	0	4	2	3	1	1	21	12089
3	11102	0	7	1	4	1	2	16	11133
4	9615	1	2	8	2	3	1	14	9646
5	7535	1	2	4	4	1	2	15	7564
6	6313	1	2	4	2	2	0	11	6335
7	5221	1	0	4	1	0	0	8	5235
8	4517	0	1	3	0	0	1	8	4530
9	4313	0	0	1	1	0	0	3	4318
10	3416	1	0	0	1	0	0	8	3426
11	2995	0	0	0	0	1	0	12	3008
12	2688	1	0	0	4	0	1	10	2704
TOTAL	95266	8	21	33	28	9	9	173	95547

Table 10. Writing Accommodations by Grade

Grade	Writing Accommodation Provided								Total
	None	Braille	Large Print	Scribe	Signing	Assistive Communication Device	Approved Nonstandard Accommodation	Not Specified	
K	12708	0	2	4	1	1	0	22	12738
1	12779	2	1	6	5	0	3	25	12821
2	12053	0	4	5	3	1	1	22	12089
3	11091	0	5	12	4	1	4	16	11133
4	9606	2	3	14	3	3	1	14	9646
5	7525	1	2	12	4	1	4	15	7564
6	6310	1	3	6	2	2	0	11	6335
7	5221	1	0	4	1	0	0	8	5235
8	4517	0	1	4	0	0	0	8	4530
9	4311	0	0	2	2	0	0	3	4318
10	3416	1	0	0	1	0	0	8	3426
11	2995	0	0	0	0	1	0	12	3008
12	2688	0	0	1	4	0	1	10	2704
TOTAL	95220	8	21	70	30	10	14	174	95547

Part 4: Test Administration

The Colorado English Language Assessment was first administered in spring 2006. In 2007 the administration was moved to winter, and the CELApro was administered to 95,547 students in January and February 2010. This test consists of four separately administered sections assessing Speaking, Listening, Reading, and Writing proficiency.

The CELApro Speaking section is individually administered. The Listening, Reading, and Writing sections may be group administered or individually administered depending upon the needs of the particular examinees being tested.

CELApro test examiners must be proficient English speakers who are able to model clear pronunciation of English phonemes. For group-administered K–2 Reading and Writing sections, students must be grouped by grade. For all of the group-administered sections, students in Grades 3 and above may be grouped either by grade or by grade span. Examiners are also instructed to group students by English proficiency in different rooms or at different times if possible.

All sections of the test are untimed in order to give students every opportunity to demonstrate their proficiency in English. The estimated administration times and administration modes are shown in Table 11. Actual times may vary.

Table 11. Estimated Administration Time and Administration Mode by Skill Area

Skill Area	Estimated Administration Time (all tests are untimed)	Administration Mode
Speaking	10 Minutes – All Grades	Individual
Listening	20 Minutes – All Grades	Group or Individual
Reading	35 Minutes – Kindergarten 45 Minutes – Grades 1–12	Group or Individual
Writing	35 Minutes – Grades K–1 45 Minutes – Grades 2–12	Group or Individual

All test examiners, school assessment coordinators (SACs), and district assessment coordinators (DACs) were instructed in standardized test administration and scoring procedures prior to the test administration.

The Speaking Subtests

The Speaking test is individually administered by a fluent English speaker who reads the test questions aloud while pointing to illustrations. All items are in constructed-response format, scored with performance-based rubrics that direct the attention of the rater (generally the examiner) to the student's use of vocabulary, social and academic language, complex grammatically correct verbal expressions, and length of responses. The Speaking test takes approximately 10 minutes per student to administer and consists of four subtests as follows:

Speak in Words
In *Speak in Words*, the examiner points to objects depicted in cue pictures and asks questions such as "What is this?" and "What is it used for?" Students respond with single words and short phrases to identify the objects and answer questions related to those objects. Student responses are scored as correct (C), incorrect (I), or no response (NR).

Speak in Sentences
In *Speak in Sentences*, students respond in complete sentences to describe activities or actions. The examiner points to each cue picture and directs the student to respond to prompts such as "Tell me what is happening in the picture," "Tell me exactly where the book is located," and "Please give me clear directions on how to go from Place A to Place B." Student responses are scored with a 0–3 rubric.

Make Conversation
Students also respond in complete sentences in *Make Conversation*. However, instead of describing pictures, students respond to the examiner's prompts such as "Tell someone to do something," "Ask someone for something," "Describe how to do something," and "Explain why we do something." Student responses are scored with a 0–3 rubric.

Tell a Story
In *Tell a Story*, students produce multiple sentences explaining what is happening in a series of four pictures. The pictures illustrate a story with a beginning, a middle, and an end. Pointing to the series of four pictures, the examiner begins the story by reading a story starter to contextualize the pictures without giving away vocabulary or key content. Student responses are scored on a 0–4 rubric.

The Listening Subtests

The Listening test is administered to a group of students by a proficient English speaker who reads from the Examiner's Guide and uses the audio CD. All Listening items are in multiple-choice format and measure general comprehension and inferential and critical thinking skills at a discourse level that integrates academic language. Students listen to classroom English to demonstrate language proficiency levels within each grade span. The Listening test takes approximately 20 minutes per group to administer and consists of three subtests: *Listen for Information, Listen in the Classroom,* and *Listen and Comprehend.*

Listen for Information
In *Listen for Information,* students hear instructions typical of those provided by a classroom teacher. Instructions vary in length from one to three sentences and must be played from the audio CD. The examiner then asks students which of three answer choices restates the instructions they heard. Instructions and answers may contain idioms and different syntactical structures.

Listen in the Classroom
Listen in the Classroom assesses comprehension of academic language, where students hear two short exchanges typical of classroom discussions. The listening passages, questions, and text answer choices must be played from the audio CD. After

listening, students respond to three questions about what they heard. Each question has three answer choices.

Grade Span	Passage Length
K–2	50–60 words
3–5	60–90 words
6–8	60–100 words
9–12	90–130 words

Listen and Comprehend
A longer listening passage included in *Listen and Comprehend* assesses comprehension of narratives. Questions ask about main ideas, details, inferences, and idioms. The listening passages, questions, and text answer choices must be played from the audio CD. Students are asked four questions about the passage. Each question has three answer choices.

Grade Span	Passage Length	Genre	Percentage
K–2	150–200 words	Fiction	83%
		Nonfiction	17%
3–5	200–250 words	Fiction	50%
		Nonfiction	50%
6–8	200–250 words	Fiction	33%
		Nonfiction	67%
9–12	225–325 words	Fiction	33%
		Nonfiction	67%

The Reading Subtests

The Reading test is usually administered to a group by a fluent English speaker who reads from the Examiner's Guide. All Reading items are in multiple-choice format. Some items evaluate phonemic awareness as the basis for recognizing words and developing vocabulary. In other items, students read literary and informational grade-appropriate texts to demonstrate sentence-level and discourse-level reading ability as well as inferential skills. The Reading test takes approximately 35–45 minutes to administer and consists of three subtests as follows:

Analyze Words

In *Analyze Words*, students respond to discrete items in a variety of formats addressing four word analysis tasks: identifying rhyming words, applying letter-sound relationships to read English words, applying letter-sound relationships to read English phonemes, and applying knowledge of morphemes and syntax to word meaning. Each question has three answer choices.

Read Words

For Grades K–2, students demonstrate vocabulary by classifying words, selecting written words to match those spoken by the examiner, and matching pictures of objects to their written descriptions. In all other grade levels, students demonstrate vocabulary by choosing synonyms or antonyms of a given word and/or choosing words that correctly complete sentences. Additionally, students in Grades 6–12 are tested on idiomatic expressions. Each question has three answer choices.

Read for Understanding

Higher-level reading skills are evaluated in *Read for Understanding*, in which students respond to passages representing various literary genres (e.g., fiction, nonfiction, and poetry). Questions address three tasks: demonstrating reading comprehension, identifying important literary features of text, and applying learning strategies to interpretation. Students in Kindergarten read along as the examiner reads passages aloud; then students identify one of three picture choices that correspond with the reading passage. Students in Grades 1 and 2 read two additional passages independently. Students in Grades 3–5 read passages without assistance and choose corresponding pictures or text. Students in upper grades read passages without assistance and select from four written answer choices.

Grade Span	Passage Length	Genre	Percentage
K	50–100 words	Fiction	100%
		Nonfiction	0%
1–2	100–150 words	Fiction	100%
		Nonfiction	0%
3–5	175–275 words	Fiction	50%
		Nonfiction	50%
6–8	250–350 words	Fiction (Poetry)	50%
		Nonfiction	50%
9–12	250–450 words	Fiction (Poetry)	50%
		Nonfiction	50%

The Writing Subtests

The Writing test is usually administered to a group by a fluent English speaker who reads from the Examiner's Guide. The test includes both multiple-choice and constructed-response items that assess both receptive and productive domains. In the first section, multiple-choice items engage students to identify appropriate grammar, mechanics, and syntax, and in the second section, students respond to prompts in the form of phrases, sentences, and paragraphs.

Responses to constructed-response items are evaluated with performance-based rubrics (on a 0–3 or 0–4 scale depending on the item) that direct the rater's attention to the student's use of English grammar and the appropriate use of discourse. The test takes approximately 35–45 minutes to administer and consists of four subtests, except that students in Grades K–1 do not take *Write in Detail*.

Use Conventions
Discrete point items in *Use Conventions* assess whether students can identify correct uses of grammar, capitalization, punctuation, and sentence structure. Each item has three answer choices.

Write About
In *Write About*, students in Grades K–1 write one sentence, and students in Grades 2–12 write two sentences to describe a picture. Responses are scored with a 0–3 rubric.

Write Why
In *Write Why*, students make a choice between two alternatives and write to explain the reason for the choice they make. In Grades K–1, students write one reason; in Grades 2–12, students write two reasons. Responses are scored with a 0–3 rubric.

Write in Detail
Prompts in *Write in Detail* elicit longer responses. Students in Grades 2 write to describe what is happening in a sequence of four pictures. Students in Grades 3–12 organize their ideas and write paragraphs or essays responding to a written prompt. Responses are scored with a 0–4 rubric. Students in Grades K–1 do not take *Write in Detail*.

Teacher Training

The Administration Training Workshops for 2010 were conducted in seven locations in Colorado: Pueblo, Colorado Springs, Durango, Grand Junction, Denver, Fort Collins, Fort Morgan, and Boulder. These locations were selected to cover the state's training needs geographically as well as in terms of district size. A total of 302 participants attended the CELApro workshops. Table 12 shows the breakdown of attendees per workshop location.

Table 12. Number of Attendees at Pre-Administration Training Workshops

Location and Date	Number of Attendees
Pueblo 11/03/09	67
Colorado Springs 11/04/09	43
Durango 11/10/09	24
Grand Junction 11/11/09	35
Denver 11/13/09	65
Fort Collins 11/16/09	15
Fort Morgan 11/17/09	30
Boulder 11/18/09	23
Total attendees	**302**

Workshop Setup

The environment of the Pre-Administration Workshop is friendly and facilitates small-group discussion. Participants' seats were not assigned. CELApro was first, followed by the CSAPA Administration Training.

Training Materials Development

The training materials were developed to reduce complexity, mirror the trainer's script, and ensure clarity in the use of the contents within the Training Folder and Training DVD throughout the training. Following are the details of the purpose of each component.

Training Folder
The CELApro Administration Training Folder contained a copy of a PowerPoint presentation, a list of acronyms, Speaking Practice Scoring Sheets, and Speaking Rubrics. The purpose of these materials was to allow for easy navigation. Navigation through the training materials is key when training a large number of participants, which in turn facilitates the learning process and helps participants gain the understanding needed to conduct their own trainings.

Training DVD
Another important part of the training materials is the coordination between the DVD and the training materials in the folder. The training DVD gives an overview of each subtest for all grade spans: Listening, Speaking, Reading, and Writing. Because the Speaking test is scored by Test Examiners during test administration, the DVD component is critical for training. CTB experts and teachers scored all speaking samples. Participants use Practice Scoring Sheets as part of the scoring calibration exercises.

Part 5: Scoring

The 2010 CELApro tests were scored and processed by CTB's scoring team using the standardized methods and procedures previously developed for the *LAS Links* program. The CELApro scoring team consists of trained technical specialists who are responsible for coordinating all scoring and reporting activities related to the processing of CELApro test documents. Document preparation, interdepartmental coordination and communication, processing specifications, and problem resolution are performed by a designated Scoring Project Manager from this team. The scoring team works closely with all CTB departments to ensure successful scoring and reporting.

CTB maintains a professional staff of specialized data processing technicians to lead the verification process and ensure the integrity of the student response data at both group and individual levels. Raw scoring and editing of scanned data is performed in a client/server system (WinScore), where a sophisticated system of edits are invoked to review the integrity of each batch scanned and to produce a list of error suspects. While the editors can view data from any document online, the error suspect list concentrates on the most likely problems based on pre-defined guidelines. This system reduces editing time and provides a high degree of quality control. CTB continues to enhance the capability of editing software to simplify the detection and correction of errors. Online editing screens focus an editor on potential problems and then provide related information. The actual scanned documents are always available to the editor, and the software supports the review and correction of any field in the scanned record. Entry and verification of the necessary corrections are enhanced to ensure each error is actually corrected. As batches are extracted for scoring, a final edit is performed to ensure all requirements for scoring are met. This automated final edit flags a batch for further editing if any error is still detected. A batch containing errors cannot be extracted for reporting. This ensures a high level of accuracy of the scored data.

When the editing process is completed, documents are moved to a staging area to be prepared for retention. Bundles are caged, warehoused in a recoverable location, and retained for possible retrieval during the specified retention period. Once this period is over, documents are destroyed according to procedures that ensure security is maintained.

Handscoring Process

For the CELApro assessments, CTB's imaging handscoring system presents images of scanned test books to trained readers, who assign scores for constructed-response items. Scanned images are viewed on high quality 19″ workstation monitors. Images of each student's responses are automatically routed to two or more readers when required, and images of specific subsets of test items are routed to designated groups of readers trained to score these items. CTB is committed to using the finest imaging equipment, software presentation system, data management system, and quality control to provide valid, reliable, cost-efficient scoring.

Readers
In order to work as a handscoring reader at CTB, one must possess and show evidence of either a BA or BS degree. The evaluator staff is comprised of individuals from many walks of life —from retired or current educators to engineers—all possessing BAs to PhDs.

Many CTB readers also have a great deal of classroom teaching experience. Our reader pool includes editors, published authors, and a number of individuals with advanced degrees. The minimum qualification for all Scoring Center readers is a Bachelor's degree.

Team Leaders
Scoring team leaders are selected on the basis of having demonstrated a high degree of scoring accuracy and consistency, often across multiple subjects and grades. They must also possess good interpersonal and leadership skills in order to be effective when training and counseling readers. The ratio of readers to team leaders is no more than 10 to 1. While it is possible to conduct handscoring with more readers per team leader, it has been CTB's experience that inter-rater reliability and production goals are jeopardized unless a trained leader can frequently monitor all readers.

Scoring Supervisors
Scoring Supervisors are the core group at CTB scoring centers. They direct and organize the assessment process, and train team leaders and readers. Scoring Supervisors have extensive experience as team leaders prior to their qualification and selection. The Scoring Supervisors are subject area experts in the content(s) that they supervise and train.

Anchor and Training Papers
Prior to the actual scoring, the CTB Scoring Center creates training materials. The process includes several presorting steps and subsequent iterative/consensus processes in order to achieve ever-increasing agreement and precision through a kind of "round robin" scoring, followed by discussion and selection. When all papers for a form are selected and assigned status as good anchors, training, qualifying, or check-set papers, they are consolidated into training formats. Scoring Guides (consisting of rubrics, anchors, and annotations) serve as a constant, setting the course for all subsequent training and scoring.

Rater Training and Validation
Validation is a critical task in the assessment training process. It is the final determinant in reader readiness. All readers, including team leaders, must achieve 80 percent exact agreement on the qualifying round following training. Those readers not validating on the first attempt receive further training prior to taking an additional qualifying round. Only those who successfully validate are qualified as readers and allowed to score tests. Team leaders are required to complete two validation rounds with 80 percent exact agreement in each round.

Intra-rater Reliability

Throughout the course of the handscoring process, calibration sets of pre-scored papers (check-sets) are administered daily to the team leaders as well as to the readers to monitor scoring accuracy and to maintain a consistent focus on the established rubric and guidelines. Imaging permits this monitoring without reader knowledge of when a check-set is administered. Readers whose check-set scores fall below the qualifying level are removed from live scoring and are given additional training and another qualifying (validation) round. Readers unable to qualify are dismissed.

The "read-behind" is another valuable intra-rater reliability monitoring technique. On a daily basis, each team leader reads a random selection of each reader's scored items. The scores are compared, and if they agree, the team leader is able to offer feedback, which enhances the reader's confidence and ability to score quickly and accurately. However, if an individual is

straying from the standard established in the training and validation samples, the aberrant scoring is detected, and the team leader is able to offer the guidance necessary to refocus the reader's effort. Readers whose scoring is inconsistent are read behind more frequently by their team leaders. Thus, any scoring variation is corrected.

Inter-rater Reliability

Intraclass correlation coefficients and weighted Kappa coefficients were calculated to measure reader agreement (Fleiss & Cohen, 1973) for each of the handscored CELApro items,[1, 2] using scores assigned to all item responses that received second reads. The intraclass correlation coefficients ranged from 0.88 to 0.96, with 83 percent of the coefficients greater than or equal to 0.90. The weighted Kappa values also were high[3] for all items, ranging from 0.75 to 0.92, indicating good agreement between the first and second readers. Inter-rater agreement statistics for all of the hand-scored items are shown in Table 13.

The percentages of discrepant ratings were higher in Kindergarten than in any of the higher grades, ranging from 8% to 12%. The percentage of discrepant ratings was 7% or less for all items in Grades 1 through 12.

---

[1] If agreement is perfect, both the intraclass correlation coefficient and Kappa will be equal to +1. If agreement is at chance levels, then both coefficients will be equal to zero.
[2] The intraclass correlation does not consider chance agreement between two raters, but the weighted Kappa does take into account chance agreement. Therefore, in general, weighted Kappa will have values equal to or smaller than the intraclass correlations.
[3] Kappa values between 0.75 and 0.91 represent good agreement beyond chance, and values below 0.40 indicate poor agreement.

Table 13. Inter-Rater Agreement for CELApro Writing Responses

Grade Span		Item	Max Score	% Perfect Agreement	% Adjacent Scores	% Special Codes	% Discrepant (>1 point)	Intraclass Correlation	Wtd Kappa
K–2	K	21	3	0.44	0.06	0.38	0.12	0.90	0.80
		22	3	0.41	0.05	0.43	0.12	0.93	0.85
		23	3	0.49	0.06	0.37	0.09	0.94	0.89
		24	3	0.48	0.06	0.38	0.08	0.94	0.88
		25	3	0.48	0.06	0.39	0.08	0.94	0.88
	1	21	3	0.68	0.19	0.09	0.04	0.91	0.83
		22	3	0.68	0.18	0.11	0.04	0.93	0.86
		23	3	0.70	0.15	0.11	0.04	0.94	0.88
		24	3	0.68	0.16	0.11	0.05	0.93	0.87
		25	3	0.70	0.14	0.11	0.05	0.93	0.87
	2	21	3	0.70	0.20	0.06	0.04	0.93	0.85
		22	3	0.70	0.21	0.06	0.02	0.92	0.84
		23	3	0.69	0.21	0.06	0.05	0.91	0.82
		24	3	0.64	0.23	0.06	0.07	0.90	0.79
		25	4	0.59	0.29	0.08	0.04	0.93	0.85
3–5		22	3	0.77	0.18	0.03	0.02	0.90	0.80
		23	3	0.77	0.18	0.03	0.02	0.91	0.81
		24	3	0.73	0.20	0.04	0.04	0.88	0.77
		25	3	0.73	0.19	0.04	0.03	0.89	0.79
		26	4	0.66	0.22	0.08	0.04	0.90	0.81
6–8		21	3	0.80	0.15	0.03	0.01	0.90	0.81
		22	3	0.82	0.14	0.03	0.01	0.91	0.82
		23	3	0.75	0.21	0.03	0.01	0.88	0.75
		24	3	0.77	0.18	0.03	0.01	0.89	0.79
		25	4	0.65	0.30	0.04	0.01	0.89	0.79
9–12		21	3	0.80	0.12	0.07	0.00	0.95	0.90
		22	3	0.81	0.11	0.07	0.00	0.96	0.92
		23	3	0.73	0.18	0.08	0.01	0.93	0.85
		24	3	0.74	0.17	0.08	0.01	0.93	0.87
		25	4	0.70	0.21	0.08	0.01	0.94	0.88

Scoring and Technology Quality Control Procedures

The Technology and Scoring Departments at CTB both have quality assurance groups specifically charged with reviewing scoring data and reports during all stages of the process. The Technology quality assurance team verifies the accuracy of all reporting programs before they become operational. The Scoring quality assurance team verifies the accuracy of report information during the scoring process. After all data are entered into the scoring system and all reporting programs are completed, a sample of reports are printed and submitted to the Scoring quality assurance group, which reviews the sample reports to verify the accuracy and correct presentation of all data.

Numerous quality assurance checks are in place throughout the scoring process to ensure the accuracy of reports. Prior to delivering any electronic files or hard-copy score reports, all reports undergo a final, extensive quality check, known as a "Red Team Review." Red Teams are comprised of individuals from every CTB department coming together to form an interdisciplinary team. Samples of each type of report are printed from the active scoring system, and the Red Team carefully reviews these samples for accuracy and correct format. Student-level information is compared by hand with student rosters and other documentation. Reports are not sent out until all necessary corrections determined by the Red Team are resolved.

Part 6: Data Analysis and Results

This section of the technical report contains a description of the 2009 calibration and equating and differential item functioning (DIF) procedures and results, along with details of the 2010 classical item analysis that was conducted for each test. Because the 2010 CELApro tests were identical to those administered in 2009, no new calibration, equating, or DIF analyses were conducted this year. The 2010 CELApro tests were scored using the raw-to-scale score tables that were produced from the results of the 2009 calibration/equating/scaling analyses. This section also includes a subsection describing student performance on the 2010 tests, along with comparisons with the 2009 and 2008 results.

IRT Item Calibration

Student item responses on each of the CELApro assessments were calibrated in 2009 using the three-parameter logistic (3PL) model to scale the multiple-choice (SR) items, and the two-parameter partial credit (2PPC) model to scale the constructed-response (CR) items. A brief explanation of the models is provided below.

The 3PL model (Lord & Novick, 1968; Lord, 1980) defines a multiple-choice item in terms of three item parameters: (a) item discrimination, (b) item difficulty or location, and (c) probability of a student with very low ability answering the item correctly (i.e., a guessing parameter). In this model, the probability that a student with scale score θ will respond correctly to item j is defined as

$$p_j(\theta) = c_j + \frac{(1 - c_j)}{1 + \exp[-1.7a_j(\theta - b_j)]},$$

where a_j is the item discrimination,
b_j is the item difficulty, and
c_j is the probability of a correct response by a very low ability student.

The 2PPC model defines a constructed-response item in terms of item discrimination as well as location parameter for each score point. The 2PPC model is a special case of Bock's (1972) nominal model. Bock's model states that the probability of an examinee with ability θ having a score at the kth level of the jth item is

$$P_{jk}(\theta) = P(x_j = k - 1 | \theta) = \frac{\exp Z_{jk}}{\sum_{i=1}^{m_j} \exp Z_{ji}}, k = 1,...,m_j,$$

where m_j is the number of score levels, and

$$Z_{jk} = A_{jk}\theta + C_{jk},$$
$$A_{jk} = \alpha_j(k-1), \ k = 1, 2,...m_j, \text{ and}$$

$$C_{jk} = -\sum_{i=0}^{k-1} \gamma_{ji}, \text{ where } \gamma_{j0} = 0,$$

where A_{jk} is the discrimination parameter of the kth category of item j, C_{jk} is the intercept parameter of the nonlinear response function associated with the kth category of item j, α_j and γ_{ji} are the parameters to be estimated from the data.

For each item there are $m_j - 1$ independent γ_{ji} parameters and one α_j parameter; a total of m_j independent item parameters is estimated.

All of the CELApro assessments were calibrated in 2009 using the 3PL/2PPC models described above. Separate calibrations were conducted for Listening, Speaking, Reading, Writing, Comprehension, and Oral scales in each grade span.

Equating and Scaling

The calibrated tests were placed on the existing CELApro/*LAS Links* scale in 2009 through a Stocking and Lord (1983) characteristic curve equating procedure. The 2008 operational item parameters for almost all of the test items were used as equating anchors in this procedure.

The 2009 M1 and M2 conversion parameters were computed as follows:

M1$_{New}$ = A* *M1*$_{Old}$

M2$_{New}$ = A* *M2*$_{Old}$ + B

where

M1$_{New}$ and *M2*$_{New}$ are the new transformation constants calculated to place the 2009 test items onto the existing scale, and

M1$_{Old}$ and *M2*$_{Old}$ are the transformation constants from the anchor set.

The *A* and *B* values are derivatives of the input (initial) and estimated (final) values for the anchor set and are computed as follows:

$$A = \frac{SD_{New}}{SD_{Old}}$$

$$B = \left(Mean_{New} - \frac{SD_{New}}{SD_{Old}} Mean_{Old} \right)$$

where

SD$_{New}$ is the standard deviation of anchor estimates in scale score metric,

SD$_{Old}$ is the standard deviation of anchor input values in scale score metric,

Mean$_{New}$ is the mean of anchor estimates in scale score metric, and

Mean$_{Old}$ is the mean of anchor input in scale score metric.

This equating procedure was performed in 2009 for each of the grade spans K, 1, 2, 3–5, 6–8, and 9–12. Consequently, the equated results were used to create raw-to-scale score tables for each of the six content areas (Reading, Writing, Listening, Speaking, Oral, and Comprehension). Because the total score is computed as the unweighted mean of the scale scores on Reading, Writing, Listening, and Speaking, no separate calibration, equating, scaling, or scoring table was required for the total score.

The resulting scoring tables for all grade spans, which were used to score the 2009 and 2010 CELApro, are included in Appendix E.

Results of the 2009 Calibration and Equating

Tables B1 to B32 and Figures B1 to B64 in Appendix B show the alignment of the original and equated "a" parameters (using the log of a) and the alignment of the corresponding "b" parameters for Listening, Speaking, Reading, and Writing. In these figures, the original parameters are the 2008 CELApro item parameters, and the equated parameters are the 2009 CELApro parameters. Since no equating was performed, the 2010 CELApro parameters are identical to the 2009 item parameters.

Figures C1 to C12 in Appendix C show the CELApro test characteristic curves (TCCs) and the conditional standard errors of measurement (CSEMs) for each grade span and content domain. For a vertically scaled test such as the CELApro/*LAS Links*, we would expect to see a pattern in which the TCCs are arrayed in grade-level sequence from left to right (i.e., with tests increasing in difficulty as grade level increases). The TCCs show this expected pattern.

The correlations between the 2009 equated and input anchor item parameters and *p*-values (*P*) are shown in Table 14. For multiple-choice scales, these represent the correlations of the *a* and *b* parameters. For constructed-response items, the correlations of item parameters represent the alpha and gamma correlations, respectively.

Table 14. Stocking and Lord Parameter Correlations

Grade Span K–2			
	P	Discrimination	Location
Speaking	1.00	0.99	1.00
Listening	1.00	0.97	0.93
Reading	1.00	0.98	1.00
Writing	0.99	0.83	0.97
Comprehension	1.00	0.98	1.00
Oral	1.00	0.93	0.95

Table 14. Stocking and Lord Parameter Correlations (continued)

Grade Span 3–5			
	P	Discrimination	Location
Speaking	1.00	0.98	0.97
Listening	1.00	0.98	1.00
Reading	1.00	0.98	0.99
Writing	0.97	0.99	1.00
Comprehension	1.00	0.98	1.00
Oral	1.00	0.96	0.99
Grade Span 6–8			
	P	Discrimination	Location
Speaking	1.00	0.98	0.98
Listening	1.00	0.94	0.98
Reading	1.00	0.98	0.98
Writing	1.00	0.96	0.98
Comprehension	1.00	0.98	0.99
Oral	1.00	0.90	0.98
Grade Span 9–12			
	P	Discrimination	Location
Speaking	1.00	0.98	0.98
Listening	1.00	0.96	0.97
Reading	1.00	0.97	0.99
Writing	1.00	0.97	0.99
Comprehension	1.00	0.97	0.99
Oral	1.00	0.98	0.99

For all contents and grade spans, the *p*-value correlations are all greater than 0.95.

For each of the six content domains, Appendix D contains the test characteristic curves for the anchor item input parameters, the equated anchor item estimated parameters, and the equated total test. As shown in these plots, the total test and the anchor test are closely aligned to each other.

Item Analysis

Classical item analysis statistics were computed for the 2010 CELApro administration for each content domain at each grade span. The tables in Appendix A present item-level descriptive statistics for each grade span and content domain. These tables contain the following information: item number, item type, item *p*-value, item correlation with the total test score, correlation between each item choice and the total test score, and percent omit. The *p*-value for an SR item represents the proportion of students who answered the item correctly. The *p*-value for a CR item represents the mean raw score for the item divided by the maximum possible score for that item.

The point biserial correlation between the item score and the total score on the test was also computed for each of the SR items. For each CR item, the Pearson product-moment correlation between the item score and the total score on the test was computed. For these correlations, the studied item was excluded from the computation of the total score so as not to artificially inflate the correlation.

Item Difficulty Statistics (*p*-values)

The 2010 statistics for individual items at each grade span are provided in the item analysis tables in Appendix A. In these tables, item difficulty is expressed in terms of *p*-values. For multiple-choice items, the *p*-value is the proportion of students answering the item correctly. For constructed-response items, the *p*-value is the mean item score expressed as a proportion of the total score points possible on that item. (i.e., each raw item score is divided by the maximum possible score on the item). The item-level results, overall, are consistent with the results obtained in 2009.

The *p*-values in Appendix A are at or above 0.20 except for eight items; and most are in the desired difficulty range between 0.30 and 0.90. The range of *p*-values varies by grade span and content domain. Across grade spans, the *p*-values range from 0.09 to 0.98 for Speaking; 0.23 to 0.98 for Listening; 0.19 to 0.99 for Reading; 0.10 to 0.94 for Writing; 0.23 to 0.99 for Comprehension; and 0.09 to 0.98 for Oral. Within grade spans, *p*-values range from 0.09 to 0.99 in Grade Span K–2; from 0.38 to 0.98 in Grade Span 3–5; from 0.40 to 0.93 in Grade Span 6–8; and from 0.36 to 0.90 in Grade Span 9–12.

Average item difficulty for each content area, grade, and grade span is summarized in Table 15. In this table, item difficulty is expressed in terms of *p*-values. For multiple-choice items, the *p*-value is the proportion of students answering the item correctly. For constructed- response items, the *p*-value is the mean item score expressed as a proportion of the total score points possible on that item (i.e., each raw item score is divided by the maximum possible score on the item).

Table 15. Mean *P*-Values by Grade Span and by Grade

Grade	Speaking	Listening	Reading	Writing	Comprehension	Oral
Grade Span 1	0.68	0.68	0.67	0.49	0.67	0.68
K	0.53	0.47	0.54	0.27	0.51	0.50
1	0.71	0.72	0.66	0.51	0.66	0.72
2	0.80	0.84	0.81	0.69	0.81	0.82
Grade Span 2	0.82	0.71	0.68	0.76	0.70	0.77
3	0.78	0.65	0.59	0.70	0.62	0.71
4	0.83	0.72	0.69	0.77	0.70	0.77
5	0.86	0.77	0.75	0.82	0.77	0.81
Grade Span 3	0.81	0.79	0.68	0.77	0.74	0.80
6	0.80	0.76	0.65	0.76	0.70	0.78
7	0.81	0.79	0.69	0.77	0.74	0.80
8	0.82	0.81	0.71	0.78	0.76	0.82
Grade Span 4	0.81	0.77	0.64	0.76	0.70	0.79
9	0.79	0.74	0.59	0.75	0.66	0.77
10	0.81	0.76	0.63	0.75	0.69	0.78
11	0.82	0.78	0.67	0.77	0.73	0.80
12	0.82	0.79	0.67	0.76	0.73	0.81

Item-Total Correlations

An important indicator of item quality is the correlation of scores on that item with scores on the total test. The 2010 item-total correlations (point biserial correlation coefficients for SR items and Pearson product-moment correlations for CR items) are summarized in Table 16. To compute these correlations, the "total" score was defined as the total score on the specific content domain. To avoid artificially inflating the correlation coefficients, the contribution of the item in question was removed from the total when calculating each of the correlations. Thus, performance on each Listening item was correlated with the total Listening score minus the score on the item in question, performance on each Speaking item was correlated with the total Speaking score minus the score on the item in question, and so on for the Reading, Writing, Oral, and Comprehension scales.

Individual item-total correlations for each content area and grade span are provided in the item analysis tables in Appendix A. Across Grades 1–12, item-total correlations for Speaking range from 0.25 to 0.85. Listening items range from 0.23 to 0.54. Item-total correlations for Reading, range from 0.16 to 0.57, and for Writing, the correlations range from 0.07 to 0.78. Comprehension item-total correlations range from 0.16 to 0.57, and Oral item-total correlations range from 0.12 to 0.82. Item-total correlations for Kindergarten were slightly lower than the other grades, ranging from 0.31 to 0.83 for Speaking, from 0.19 to 0.53 for Listening, from -0.01

to 0.48 for Reading, from 0.07 to 0.61 for Writing, from -0.03 to 0.49 for Comprehension, and from 0.13 to 0.80 for Oral.

The average (mean) item-total correlation coefficients for each content area, grade span, and grade are shown in Table 16. The average item-total correlation coefficients ranged from 0.54 to 0.67 for Speaking, from 0.34 to 0.43 for Listening, 0.35 to 0.44 for Reading, 0.26 to 0.49 for Writing, 0.33 to 0.42 for Comprehension, and 0.38 to 0.49 for Oral.

Table 16. Average Item-Total Correlations by Grade Span and Grade

Grade	Speaking	Listening	Reading	Writing	Comprehension	Oral
Grade Span 1	0.60	0.39	0.36	0.39	0.34	0.44
K	0.63	0.38	0.36	0.26	0.34	0.44
1	0.60	0.41	0.35	0.43	0.33	0.44
2	0.58	0.39	0.39	0.49	0.36	0.43
Grade Span 2	0.55	0.36	0.43	0.48	0.37	0.40
3	0.54	0.34	0.42	0.48	0.35	0.38
4	0.55	0.36	0.43	0.48	0.37	0.40
5	0.56	0.37	0.44	0.48	0.39	0.42
Grade Span 3	0.61	0.40	0.40	0.43	0.38	0.44
6	0.57	0.38	0.39	0.43	0.36	0.42
7	0.61	0.40	0.40	0.43	0.39	0.45
8	0.63	0.42	0.42	0.44	0.40	0.47
Grade Span 4	0.66	0.42	0.41	0.48	0.41	0.49
9	0.67	0.41	0.39	0.48	0.39	0.48
10	0.67	0.42	0.41	0.49	0.40	0.49
11	0.65	0.43	0.41	0.47	0.41	0.49
12	0.65	0.43	0.43	0.48	0.42	0.49

Item Omit Rates

The item analysis tables in Appendix A also show the rate at which students omitted items. Omit rates are often useful in determining whether testing times are sufficient, particularly if there is a high rate of items omitted at the end of a test section. In cases where speededness is not an issue, high item omit rates may often indicate ambiguity or extreme item difficulty.

Omit rates were generally low for students in Grades 3 through 12. Omit rates for Grade Spans 6–8 and 9–12 were below 5 percent for all of the items in all content areas. For Grade Span 3–5, Writing had one item with an omit rate of 8.27 percent and two Reading items had omit rates of 14.63 percent and 15.42 percent.

Omit rates were generally higher for Grade Span K–2. Omit rates were between 0.27 and 8.49 percent for all of the Listening items, with nine items above 5 percent. For the Reading items,

omit rates were above 5 percent for 12 items in Grades 1–2 and for 21 items administered to Kindergarten students. Speaking K–2 had five items above 5 percent, all of them in Kindergarten. The highest omit rates were for the Kindergarten Writing items; with omit rates ranging from 2.58 percent to 23.47 percent and with all but two items above 5 percent. Writing omit rates were lower in Grade 1, and all were below 5 percent for students in Grade 2.

Differential Item Functioning (DIF) Statistics

In addition to the analyses that were conducted as part of the *LAS Links* development process, Linn-Harnisch (1981) gender DIF analyses were conducted on data from the 2009 CELApro administration. The procedures employed for CELApro DIF analyses apply IRT-based models to analyze the item calibration data. Because no new calibrations were conducted in 2010, t no new DIF analyses were conducted. The DIF analyses and results in this technical report are based on the 2009 item parameters and thus, are identical to those that were reported last year.

For the 2009 CELApro analyses, a separate IRT calibration and a separate DIF analysis were conducted for each grade span and content domain (Listening, Speaking, Reading, Writing, Oral, and Comprehension). To calculate DIF for the CELApro assessments, the IRT parameters for each item (a_i, b_i, c_i) and the trait or ability estimate (θ_i) for each examinee were estimated for the three-parameter logistic model:

$$P_{ij} = c_i + \frac{1 - c_i}{1 + \exp\left[-1.7 a_i \left(\theta_j - b_i\right)\right]},$$

where P_{ij} is the probability that examinee j will pass item i. The total population is then divided into two groups by gender, and the members in each group are sorted into ten equal score categories (deciles) based upon their location on the scale score (θ_i) scale. The expected proportion correct for each group based on the model prediction is compared to the observed (actual) proportion correct obtained by the group. The proportion of examinees in decile g who are expected to answer item i correctly is:

$$P_{ig} = \frac{1}{n_g} \sum_{j \varepsilon g} P_{ij},$$

where n_g is the number of examinees in decile g. The proportion of examinees expected to answer item i correctly (over all deciles) for a group (e.g., female) is:

$$P_i = \frac{\sum_{g=1}^{10} n_g P_{ig}}{\sum_{g=1}^{10} n_g}.$$

The corresponding observed proportion correct for examinees in a decile (O_{ig}) is defined as the number of examinees in decile g who answered item i correctly divided by the total number of examinees in the decile (n_g). That is,

$$O_{ig} = \frac{\sum\limits_{j \varepsilon g} u_{ij}}{n_g},$$

where u_{ij} is the dichotomous score for item i for examinee j.

The corresponding formula to compute the observed proportion answering each item correctly (over all deciles) for a complete gender group is given by:

$$O_i = \frac{\sum\limits_{g=1}^{10} n_g O_{ig}}{\sum\limits_{g=1}^{10} n_g}.$$

After the values are calculated for these variables, the difference between the observed proportion correct for a gender group and expected proportion correct can be computed. The decile group difference (D_{ig}) for observed and expected proportion correctly answering item i in decile g is:

$$D_{ig} = O_{ig} - P_{ig},$$

and the overall group difference (D_i) between observed and expected proportion correct for item i in the complete group (over all deciles) is:

$$D_i = O_i - P_i.$$

DIF is defined in terms of the decile group and total target subsample differences, the D_{i-} (sum of the negative group differences) and D_{i+} (sum of the positive group differences) values, and the corresponding standardized difference (Z_i) for the subsample (see Linn & Harnisch, 1981, p. 112). Items for which $|D_i| \geq 0.10$ and $|Z_i| \geq 2.58$ are flagged as DIF items. If D_i is positive, the item favors the target subsample. If D_i is negative, the item favors the standard sample.

These indices are indicators of the degree to which members of a gender group perform better or worse than expected on each item, based on the parameter estimates from all subsamples. Differences for decile groups provide an index for each of the ten regions on the scale score (θ) scale. The decile group difference (D_{ig}) can be either positive or negative. Use of the decile group differences as well as the overall group difference allows one to detect items that give a

large positive difference in one range of θ and a large negative difference in another range of θ, yet have a small overall difference. A generalization of the Linn and Harnisch (1981) procedure was used to measure DIF for constructed-response items.

The results of the 2009 DIF analyses are shown in Table 17. Again, since calibrations did not occur, the results are the same as last year. Overall, very few items exhibited differential item functioning by ethnicity. Across all grades and content areas, no item was flagged for DIF against males or females. Across all grades and content areas, 24 items (3.02%) were flagged in favor of and 29 items (3.65%) were flagged against American Indian/Alaska Natives, 11 items (1.39%) were flagged in favor of and 13 (1.64%) items were flagged against Asian/Pacific Islanders, 27 (3.40%) items were flagged in favor of and 23 items (2.90%) were flagged against Black examinees, and 2 items (0.25%) were flagged against White examinees.

All items flagged for DIF are carefully reviewed by CTB's content development experts to try to determine whether race, native language, or another characteristic might have caused the DIF. If that review suggests that the DIF statistics are likely to reflect racial bias rather than only meaningful language differences, the items will be replaced in revised future forms whenever suitable replacement items are available.

Table 17. Number of Items Exhibiting Differential Item Functioning

Subject	Grade Span	American Indian/Alaska Native		Asian/Pacific Islander		Black		White	
		For	Against	For	Against	For	Against	For	Against
Speaking	K–2	0	2	1	0	0	0	0	0
	3–5	1	1	0	0	0	1	0	0
	6–8	2	3	0	1	3	2	0	0
	9–12	0	2	1	0	3	4	0	0
Listening	K–2	2	1	0	0	0	0	0	0
	3–5	0	1	0	0	0	0	0	0
	6–8	0	0	0	0	0	0	0	0
	9–12	0	0	0	0	1	1	0	0
Reading	K–2	0	2	0	0	0	0	0	0
	3–5	1	2	1	0	0	0	0	0
	6–8	2	1	0	0	1	1	0	0
	9–12	1	2	1	4	3	2	0	1
Writing	K–2	0	0	1	0	0	0	0	0
	3–5	0	0	0	1	0	0	0	0
	6–8	1	2	1	0	0	0	0	0
	9–12	1	0	3	2	0	2	0	0
Comprehension	K–2	2	1	0	0	0	0	0	0
	3–5	0	2	0	0	0	0	0	0
	6–8	0	0	0	0	1	0	0	0
	9–12	0	2	1	4	4	3	0	1
Oral	K–2	4	1	1	0	2	0	0	0
	3–5	1	2	0	0	1	1	0	0
	6–8	2	1	0	1	5	2	0	0
	9–12	4	1	0	0	3	4	0	0

Student Performance on the 2010 CELApro

This section of the report summarizes the performance of students on the 2010 CELApro. Results are presented for the total population and for various subgroups of interest. In addition, results are compared with performance on the 2009 CELApro. To facilitate interpretation of the score distributions provided in this report, the lowest obtainable scale scores (LOSS) and the highest obtainable scale scores (HOSS) on the CELApro are provided in Table 18. These values do not change from year to year.

Table 18. CELApro Lowest and Highest Obtainable Scale Scores

		Speaking	Listening	Reading	Writing	Comp (R+L)	Oral (L+S)	Total
Grade K	LOSS	300	300	240	200	270	280	260
	HOSS	580	560	570	630	570	620	585
Grade 1	LOSS	300	300	240	200	270	280	260
	HOSS	580	560	590	630	590	620	590
Grade 2	LOSS	300	300	240	200	270	280	260
	HOSS	580	560	590	640	590	620	592
Grades 3–5	LOSS	310	310	300	270	320	290	297
	HOSS	635	630	660	680	660	680	651
Grades 6–8	LOSS	325	360	380	300	360	310	341
	HOSS	645	640	690	690	680	700	666
Grades 9–12	LOSS	330	370	390	310	380	320	350
	HOSS	650	650	700	700	700	710	675

Note: LOSS = Lowest Obtainable Scale Score; HOSS = Highest Obtainable Scale Score

Table 19 shows the 2010 total scale score means and standard deviations by grade span, and Table 20 shows the results for each individual grade in 2008, 2009, and 2010.

Table 19. 2010 Total Scale Score Means and Standard Deviations by Grade Span.

	N	Mean	SD
Grade Span 1	36471	440.81	55.48
Grade Span 2	27823	525.11	44.12
Grade Span 3	15698	553.54	42.85
Grade Span 4	12278	548.41	45.53

Table 20. 2008, 2009, and 2010 Total Scale Score Means and Standard Deviations by Grade

	2008			2009			2010		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
K	10198	383.48	37.65	11196	384.57	37.86	12073	388.87	38.52
1	11844	441.78	40.94	12242	443.84	41.47	12539	447.63	39.64
2	10714	481.55	37.97	11592	484.01	38.16	11859	486.48	37.60
3	8845	498.14	39.93	10058	500.10	40.01	10926	508.02	41.09
4	7573	519.72	41.55	7961	521.46	41.06	9473	529.19	41.50
5	6444	536.80	41.05	6900	537.73	42.00	7424	545.08	41.97
6	5189	537.22	41.64	5434	539.77	41.98	6198	546.65	40.61
7	4485	544.78	46.06	4660	546.33	44.84	5100	555.04	42.49
8	3812	550.50	46.97	3855	553.09	47.37	4400	561.50	44.74
9	3701	535.31	45.20	3674	534.27	46.64	4056	541.33	44.27
10	2963	545.56	44.18	3176	543.85	47.34	3174	547.03	46.06
11	2344	548.88	44.88	2511	551.40	45.51	2737	554.90	44.45
12	1742	549.35	44.29	1993	550.31	46.68	2311	555.08	46.23

The 2010 total scale scores were higher than the 2008 and 2009 scores in all grades. The greatest increase in scores occurred in Grades 7 and 8.

The 2010 performance on the six component scales of Speaking, Listening, Reading, Writing, Comprehension, and Oral Proficiency is summarized by grade and by grade span in Table 21 and by grade and gender in Table 22. Note that because not all students had valid scores on all four components, the component scale score means may be based upon larger numbers of students than the total scale score means.

Overall, female students tended to score somewhat higher than male students. The greatest gender differences were observed in Reading and Writing. Female students scored higher than male students on the Writing test at all grade levels. Differences in the mean Writing scores were most evident in the elementary school years where the female score advantage ranged from 9 points to more than 18 points, with smaller differences observed at higher grade levels. Male students, on the other hand, tended to score substantially higher than females on the Speaking test in Grades 8 through 12. The difference in mean Speaking scores was highest in Grades 9 and 11, where the mean score for male students was nearly 13 points higher than the mean for female students. These results are displayed graphically in Figures 1 through 7.

Table 21. CELApro Scale Score Means and Standard Deviations: Component Scales

	Speaking			Listening			Reading			Writing			Comprehension			Oral		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Grade Span 1	37277	486.44	44.26	37158	463.31	47.54	37092	412.55	65.65	36675	399.26	97.28	37021	443.27	54.19	37046	476.62	41.90
K	12605	461.70	43.28	12516	425.72	38.35	12483	355.86	46.96	12130	310.09	78.22	12448	396.89	41.08	12485	448.77	42.60
1	12704	490.54	38.53	12662	468.10	37.48	12652	418.31	46.36	12598	413.50	71.02	12627	447.60	38.73	12636	481.26	32.16
2	11968	508.12	37.69	11980	497.50	36.30	11957	465.65	50.84	11947	474.77	59.45	11946	487.01	39.79	11925	500.85	32.19
Grade Span 2	28074	540.57	47.34	28021	515.39	48.09	28018	515.46	55.70	28009	529.42	61.75	27954	514.43	46.91	27927	532.04	42.98
3	11040	528.73	44.99	10997	497.80	43.61	11022	494.48	52.49	11012	511.20	59.81	10973	495.94	42.83	10965	517.59	37.73
4	9535	543.03	46.28	9544	519.14	45.79	9521	521.12	52.13	9530	534.30	59.47	9518	518.61	43.46	9502	534.61	41.11
5	7499	554.87	47.64	7480	536.46	47.73	7475	539.19	53.40	7467	550.06	59.75	7463	536.29	46.21	7460	549.99	45.09
Grade Span 3	15828	556.79	54.34	15809	555.40	56.81	15802	545.98	47.72	15785	556.28	53.56	15778	542.07	43.92	15740	552.37	49.99
6	6243	551.02	51.32	6231	546.04	54.45	6230	537.37	44.73	6225	552.44	53.54	6221	533.77	41.09	6211	544.86	45.40
7	5145	558.22	54.46	5144	557.03	56.54	5139	547.82	46.77	5137	557.50	53.29	5131	543.97	43.28	5118	553.67	49.83
8	4440	563.26	57.40	4434	566.68	58.14	4433	555.97	50.63	4423	560.28	53.55	4426	551.53	46.29	4411	561.45	54.50
Grade Span 4	12546	547.03	59.91	12440	547.69	59.45	12532	553.58	45.00	12507	545.32	53.37	12403	551.48	50.39	12328	546.11	56.75
9	4120	541.39	59.36	4091	539.25	56.85	4111	543.75	43.71	4100	541.50	53.20	4082	541.47	47.69	4071	538.04	53.27
10	3234	547.02	60.01	3205	545.80	59.72	3233	551.32	45.20	3232	544.00	54.10	3199	549.11	49.96	3185	545.37	56.93
11	2809	551.32	59.59	2773	555.13	59.55	2802	562.54	42.57	2797	550.29	51.45	2763	560.40	49.73	2748	552.30	57.79
12	2383	551.74	60.29	2371	556.08	61.16	2386	563.05	45.75	2378	547.86	54.31	2359	561.57	52.55	2324	553.92	59.23

Table 22. CELApro Scale Score Means and Standard Deviations by Grade and Gender

		Speaking			Listening			Reading			Writing			Comprehension			Oral		
		N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
K	F	6082	462.93	44.17	6041	427.45	38.07	6023	358.46	46.31	5852	314.98	78.37	6007	399.26	40.80	6027	449.89	43.35
	M	6520	460.54	42.42	6472	424.09	38.56	6457	353.42	47.43	6275	305.57	77.82	6438	394.67	41.22	6455	447.72	41.86
1	F	6200	492.20	39.80	6182	469.38	36.91	6180	420.87	45.71	6154	420.28	68.29	6167	449.47	38.21	6174	482.55	33.17
	M	6502	488.95	37.19	6478	466.87	37.96	6470	415.84	46.84	6442	407.00	72.96	6458	445.79	39.14	6460	480.01	31.12
2	F	5769	510.05	37.67	5776	498.60	35.35	5760	468.54	49.54	5765	481.39	55.78	5755	489.06	38.56	5747	502.41	32.02
	M	6198	506.33	37.63	6203	496.48	37.15	6196	462.96	51.88	6181	468.59	62.05	6190	485.11	40.81	6177	499.40	32.29
3	F	5366	529.00	44.67	5347	496.47	42.15	5358	497.90	51.75	5352	519.81	57.47	5333	497.41	41.35	5334	517.16	36.89
	M	5672	528.48	45.30	5648	499.04	44.91	5662	491.26	52.98	5658	503.08	60.84	5638	494.54	44.16	5629	517.98	38.50
4	F	4643	542.43	46.53	4645	517.69	44.56	4637	524.37	50.33	4640	542.20	58.00	4637	519.69	42.30	4627	533.57	40.86
	M	4890	543.60	46.05	4897	520.49	46.89	4882	518.00	53.59	4888	526.77	59.88	4879	517.57	44.50	4873	535.59	41.34
5	F	3505	555.24	46.50	3498	535.90	46.08	3500	542.70	51.62	3496	559.84	58.46	3491	538.12	44.28	3488	549.80	43.85
	M	3994	554.54	48.63	3982	536.95	49.13	3975	536.09	54.74	3971	541.45	59.55	3972	534.68	47.79	3972	550.16	46.16
6	F	2946	550.69	51.28	2942	549.22	54.08	2937	541.37	43.57	2935	558.43	52.56	2935	537.00	39.93	2932	545.84	45.55
	M	3296	551.31	51.37	3288	543.18	54.64	3292	533.81	45.45	3289	547.10	53.84	3285	530.89	41.90	3278	543.98	45.26
7	F	2395	557.64	52.25	2397	561.47	55.68	2393	552.30	46.20	2395	564.24	51.09	2390	547.98	43.01	2383	554.69	48.56
	M	2748	558.71	56.33	2745	553.19	57.01	2744	543.89	46.92	2740	551.60	54.46	2739	540.47	43.23	2733	552.79	50.92
8	F	2011	560.92	56.91	2008	571.05	57.16	2009	559.45	48.54	2008	565.38	52.21	2007	554.65	45.32	2000	561.49	54.63
	M	2426	565.17	57.76	2422	562.99	58.73	2420	553.04	52.17	2412	555.94	54.25	2415	548.89	46.97	2408	561.39	54.43
9	F	1789	534.13	56.72	1770	537.95	56.54	1781	543.12	43.99	1778	544.68	53.25	1769	540.55	48.25	1767	533.57	51.51
	M	2330	547.01	60.72	2320	540.32	56.98	2329	544.23	43.51	2321	539.11	53.01	2312	542.22	47.22	2303	541.53	54.27
10	F	1486	540.75	58.42	1476	544.96	58.00	1490	551.20	44.17	1487	547.06	55.23	1475	548.49	48.65	1463	540.98	54.90
	M	1748	552.35	60.83	1729	546.52	61.16	1743	551.42	46.09	1745	541.39	53.00	1724	549.64	51.07	1722	549.11	58.35
11	F	1279	544.37	56.71	1256	554.95	58.85	1276	563.38	41.44	1275	552.51	50.69	1252	560.76	48.94	1246	548.42	56.20
	M	1529	557.15	61.32	1515	555.41	59.99	1523	561.81	43.49	1519	548.40	52.02	1509	560.17	50.39	1501	555.53	58.92
12	F	1134	545.97	56.89	1129	557.88	60.76	1140	564.99	45.61	1139	552.06	54.14	1124	563.54	52.24	1105	550.73	57.62
	M	1248	556.92	62.75	1241	554.48	61.51	1245	561.24	45.83	5852	314.98	78.37	1234	559.76	52.79	1218	556.80	60.56

Figure 1. Mean Speaking Scale Scores by Grade and Gender

Figure 2. Mean Listening Scale Scores by Grade and Gender

Figure 3. Mean Reading Scale Scores by Grade and Gender

Figure 4. Mean Writing Scale Scores by Grade and Gender

Figure 5. Mean Comprehension Scale Scores by Grade and Gender

Figure 6. Mean Oral Scale Scores by Grade and Gender

Figure 7. Mean Total Scale Scores by Grade and Gender

The performance of students tested with and without accommodations is provided in Tables 23 and 24. Because the numbers of students receiving accommodations at each grade level are very small, all accommodations for a content domain are combined in these tables.

Table 23. Total Scale Score Means by Grade and Accommodations

| Grade | Total Scale Score | | | | | |
| | Without Accommodations | | | With Accommodations | | |
	N	Mean	SD	N	Mean	SD
K	12064	388.87	38.53	9	385.89	32.22
1	12527	447.67	39.62	12	410.25	45.91
2	11845	486.55	37.52	14	429.21	58.49
3	10900	508.14	40.99	26	457.50	53.67
4	9448	529.27	41.47	25	497.12	42.65
5	7403	545.24	41.81	21	487.19	58.09
6	6184	546.73	40.56	14	510.93	47.21
7	5092	555.12	42.43	8	502.25	51.71
8	4394	561.53	44.76	6	535.33	21.77
9	4054	541.35	44.27	2	492.00	28.28
10	3171	547.00	46.08	3	570.67	8.08
11	2736	554.93	44.43	1	475.00	n/a
12	2307	555.21	46.12	4	482.50	59.90

Table 24. Component Scale Score Means by Grade and Accommodations

Grade	Speaking Scale Scores						Listening Scale Scores					
	No Speaking Accommodations			With Speaking Accommodations			No Listening Accommodations			With Listening Accommodations		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
K	12602	461.71	43.28	3	426.67	26.50	12513	425.72	38.36	3	426.00	3.46
1	12700	490.56	38.51	4	435.25	51.55	12657	468.11	37.47	5	437.60	35.08
2	11959	508.15	37.69	9	473.22	32.20	11972	497.53	36.30	8	458.88	25.37
3	11029	528.79	44.93	11	467.64	66.06	10982	497.86	43.52	15	454.13	78.38
4	9528	543.06	46.27	7	513.00	54.12	9537	519.15	45.78	7	499.00	59.39
5	7492	554.93	47.59	7	480.43	47.42	7472	536.55	47.65	8	452.75	47.91
6	6238	551.06	51.31	5	500.00	53.64	6224	546.06	54.44	7	528.71	67.33
7	5141	558.27	54.44	4	492.00	42.06	5141	557.10	56.48	3	443.33	52.39
8	4437	563.27	57.41	3	539.33	50.85	4431	566.71	58.14	3	517.67	47.43
9	4119	541.39	59.37	1	528.00	n/a	4090	539.27	56.84	1	439.00	n/a
10	3232	547.01	60.02	2	566.00	11.31	3202	545.74	59.71	3	606.67	34.85
11	2808	551.34	59.59	1	500.00	n/a	2772	555.18	59.51	1	424.00	n/a
12	2379	551.77	60.25	4	534.50	89.34	2366	556.27	61.05	5	470.20	56.65

Grade	Reading Scale Scores						Writing Scale Scores					
	No Reading Accommodations			With Reading Accommodations			No Writing Accommodations			With Writing Accommodations		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
K	12477	355.87	46.97	6	353.00	32.66	12123	310.07	78.21	7	359.71	87.93
1	12644	418.35	46.30	8	344.63	74.99	12585	413.54	70.97	13	372.62	105.81
2	11947	465.69	50.81	10	419.50	66.98	11934	474.83	59.40	13	420.92	78.42
3	11007	494.59	52.37	15	415.87	79.74	10986	511.34	59.72	26	453.81	68.66
4	9505	521.18	52.14	16	485.31	35.52	9505	534.41	59.45	25	491.04	53.39
5	7462	539.34	53.21	13	450.85	85.74	7445	550.32	59.44	22	462.55	94.39
6	6222	537.45	44.67	8	477.63	50.05	6214	552.48	53.50	11	527.73	70.88
7	5133	547.86	46.74	6	507.67	59.56	5131	557.53	53.30	6	536.67	43.86
8	4430	555.98	50.65	3	530.00	10.39	4419	560.30	53.57	4	531.50	21.32
9	4111	543.75	43.71	0	n/a	n/a	4098	541.53	53.19	2	479.50	50.20
10	3231	551.29	45.21	2	592.00	16.97	3230	544.01	54.11	2	521.00	53.74
11	2801	562.56	42.57	1	520.00	n/a	2796	550.32	51.43	1	457.00	n/a
12	2380	563.15	45.71	6	520.33	48.59	2373	548.07	54.06	5	447.80	84.84

Part 7: Reliability and Validity Evidence

Validity and reliability statistics were computed using the data from the Spring 2010 CELApro administration, along with information obtained in previous years. Overall, the 2010 CELApro analyses yielded results that are consistent with previous CELApro and *LAS Links* results. Test validation is an ongoing process of gathering evidence from many sources to evaluate the soundness of the desired score interpretation or use. This evidence is acquired from studies of the content of the test as well as from studies involving scores produced by the test. Additionally, reliability is a necessary element for validity. A test cannot be valid if it is not also reliable. All test scores contain some measurement error. Test score reliability refers to the degree to which scores on a particular assessment are free of the kinds of measurement errors that introduce variability in a student's scores. Thus, the reliability coefficient quantifies the expected consistency of student performance across multiple test forms or multiple testing occasions.

Internal Consistency Reliability

Total test reliability measures, such as Cronbach's coefficient alpha (1951) and standard error of measurement, consider the consistency (reliability) of performance over all test questions in a given form; the results of which imply how well the questions measure the content domain and could continue to do so over repeated administrations. Total test reliability coefficients such as coefficient alpha may range from 0.00 to 1.00, where 1.00 refers to a perfectly consistent test.

The internal consistency reliability of the CELApro Speaking, Listening, Reading, Writing, Oral, and Comprehension scales was evaluated using Cronbach's coefficient alpha, computed with the standard formula

$$C_\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_i^2}{\sigma_X^2}\right],$$

where

n = the number of items,

$\sigma^2{}_i$ = the raw item variance, and

$\sigma^2{}_X$ = the raw score variance for each scale.

Because the CELApro total scale score is a composite (the unweighted mean of the four component scale scores on Reading, Writing, Listening, and Speaking), the internal consistency reliability of the total score was computed using the following formula for the reliability of battery composites:

$$\rho_{ZZ'} = 1 - \frac{\sum\limits_{j=1}^{k} \sigma^2_{x_j}(1 - \rho_{x_j x'_j})}{k^2 \sigma^2_Z},$$

where

k = the number of component scales (for CELApro, $k=4$),

$\rho_{x_j x'_j}$ = reliability of each of the component scales,

$\sigma^2_{x_j}$ = scale score variance of each of the component scales, and

σ^2_Z = variance of the total (mean) scale score.

The internal consistency reliability coefficients for the 2010 CELApro tests are shown in Table 25. Achievement tests are typically considered to be of sound reliability when their reliability coefficients are in the range of 0.80 and above. All of the reliability coefficients for Speaking, Reading, Writing, Oral, and Comprehension meet or exceed this criterion. However, the reliability coefficients for the Listening scale are at or below 0.80 for every grade and almost every grade span. Because the Listening scores account for one fourth of the total composite, their lower reliability serves to lower the total score reliability as well. In spite of this, the total score reliability coefficients exceed 0.90 for every grade and grade span.

Table 25. Internal Consistency Reliability Coefficients by Grade Span and Grade

	Speaking	Listening	Reading	Writing	Comprehension	Oral	Total Score
Grade Span 1	0.93	0.86	0.88	0.83	0.91	0.94	0.95
K	0.93	0.75	0.83	0.81	0.84	0.92	0.91
1	0.91	0.77	0.82	0.86	0.84	0.91	0.94
2	0.90	0.72	0.86	0.88	0.86	0.90	0.94
Grade Span 2	0.88	0.70	0.88	0.88	0.88	0.89	0.94
3	0.88	0.65	0.86	0.87	0.85	0.87	0.93
4	0.88	0.68	0.87	0.87	0.86	0.88	0.94
5	0.89	0.70	0.88	0.87	0.88	0.89	0.94
Grade Span 3	0.92	0.74	0.84	0.83	0.88	0.91	0.93
6	0.90	0.72	0.83	0.83	0.86	0.90	0.93
7	0.92	0.74	0.84	0.83	0.88	0.92	0.93
8	0.93	0.76	0.86	0.84	0.89	0.92	0.94
Grade Span 4	0.94	0.78	0.85	0.87	0.89	0.93	0.95
9	0.94	0.76	0.83	0.86	0.88	0.93	0.95
10	0.94	0.78	0.85	0.87	0.89	0.93	0.95
11	0.93	0.79	0.85	0.86	0.89	0.93	0.95
12	0.93	0.80	0.86	0.88	0.90	0.93	0.95

Standard Errors of Measurement

Another measure of reliability is a direct estimate of the degree of measurement error in students' reported scores on a test. This second measure of reliability is called the standard error of measurement (SEM) and represents the number of score points about which a given score is expected to vary. The smaller the SEM, the smaller the variability and the higher the reliability. The SEM of the CELApro Speaking, Listening, Reading, Writing, Oral, and Comprehension scales was computed with the standard formula

$$SEM = SD * \sqrt{1 - alpha} ,$$

where

SD = standard deviation of scale score
alpha = reliability coefficient

The SEMs for the Spring 2010 CELApro assessments are shown in Table 26.

Table 26. Standard Errors of Measurement by Grade Span and Grade.

	Speaking	Listening	Reading	Writing	Comprehension	Oral	Total Score
Grade Span 1	11.66	17.66	22.70	39.76	16.46	10.37	12.61
K	11.63	19.11	19.46	34.30	16.38	12.07	11.34
1	11.41	18.10	19.47	26.75	15.34	9.53	9.85
2	12.04	19.31	19.11	20.50	14.83	10.23	9.02
Grade Span 2	16.12	26.13	19.02	21.53	16.33	14.54	10.51
3	15.65	25.91	19.63	21.59	16.68	13.46	10.51
4	16.05	25.94	18.69	21.26	16.02	14.36	10.40
5	16.09	26.11	18.52	21.41	16.12	15.26	10.43
Grade Span 3	15.75	28.81	18.83	22.04	15.47	14.78	10.95
6	15.90	28.88	18.55	22.05	15.37	14.39	10.95
7	15.50	28.73	18.62	22.21	15.28	14.49	10.91
8	15.58	28.19	19.15	21.73	15.54	15.08	10.83
Grade Span 4	15.05	27.83	17.34	19.41	16.47	14.88	10.24
9	14.46	27.58	17.83	19.68	16.55	13.91	10.23
10	14.98	28.08	17.53	19.59	16.48	14.88	10.32
11	15.45	27.56	16.62	18.98	16.18	15.30	10.11
12	15.74	27.67	17.00	19.08	16.34	15.78	10.21

Validity Evidence

The purpose of test validation is to validate interpretations of the test scores for particular purposes or uses. Test validation is an ongoing process, beginning at initial conceptualization and continuing throughout the lifetime of an assessment. Every aspect of an assessment provides evidence in support of its validity (or evidence to the contrary), including design, content requirements, item development, and psychometric quality.

The *LAS Links* and CELApro tests were designed and developed to provide English language proficiency scores that are valid for most types of educational decision making. The primary inferences from the test results include measurement of the proficiency of individual students relative to an international sample and relative program effectiveness based on the results of groups of students. Progress can be tracked over years and grades. The results can be used in a norm- and/or criterion-referenced manner to analyze the strengths and weaknesses of a

student's growth in each skill area, to plan for further instruction and curriculum development, and to report progress to parents. The results can also be used as one factor in making administrative decisions about program effectiveness, class grouping, needs assessment, and placement in ELD programs.

The *LAS Links* program was developed in accordance with the criteria for test development, administration, and use described in the Standards for Educational and Psychological Testing (1999) adopted by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME).

Content Validity

Content-related validity for language proficiency tests is evidenced by a correspondence between test content and instructional content. To ensure such correspondence, developers conducted a comprehensive curriculum review and met with educational experts to determine common educational goals and the knowledge and skills emphasized in curricula across the country. This information guided all phases of the design and development of the *LAS Links* suite of assessments.

As described in Part 1 of this report and summarized previously in Table 2, a study of the alignment of the CELApro assessments to the Colorado standards was also conducted, and a high level of agreement has been found. This alignment is expected to become even stronger as the CELApro assessments are further customized in future years.

Construct Validity

Construct validity, what test scores mean and what kinds of inferences they support, is the central concept underlying the *LAS Links* test validation process. Evidence for construct validity is comprehensive and integrates evidence from both content- and criterion-related validity. To establish meaningfulness, *LAS Links* should correlate highly with independent measures of achievement and cognitive ability.

Convergent and discriminate validity evidence can also be established through a pattern of high correlations among scales that purport to measure domains that are known to be closely related and lower correlations among scales that purport to measure dissimilar domains. This kind of pattern provides evidence that the scales are actually measuring the constructs that they purport to measure. While we have no external measures available at present to correlate with the CELApro scale scores, the pattern of correlations within CELApro provides preliminary validity evidence. The 2010 intercorrelations among the CELApro scales for each grade and grade span are shown in Tables 27 through 30.

Table 27. CELApro Scale Score Correlations, Grade Span K–2

		Listening	Reading	Writing	Comprehension	Oral	Total
K	Speaking	0.52	0.54	0.27	0.59	0.91	0.69
	Listening		0.63	0.28	0.88	0.71	0.71
	Reading			0.44	0.87	0.60	0.82
	Writing				0.36	0.29	0.78
	Comprehension					0.73	0.81
	Oral						0.74
1	Speaking	0.53	0.53	0.50	0.57	0.92	0.73
	Listening		0.61	0.53	0.84	0.75	0.77
	Reading			0.69	0.90	0.61	0.86
	Writing				0.66	0.56	0.89
	Comprehension					0.73	0.89
	Oral						0.82
2	Speaking	0.48	0.52	0.52	0.55	0.91	0.73
	Listening		0.55	0.51	0.78	0.71	0.73
	Reading			0.74	0.90	0.58	0.88
	Writing				0.72	0.56	0.89
	Comprehension					0.70	0.90
	Oral						0.80

Table 28. CELApro Scale Score Correlations, Grade Span 3–5

		Listening	Reading	Writing	Comprehension	Oral	Total
3	Speaking	0.49	0.49	0.49	0.54	0.91	0.73
	Listening		0.59	0.55	0.83	0.77	0.78
	Reading			0.73	0.89	0.60	0.87
	Writing				0.73	0.58	0.87
	Comprehension					0.74	0.91
	Oral						0.85
4	Speaking	0.48	0.49	0.47	0.53	0.89	0.73
	Listening		0.61	0.54	0.84	0.79	0.79
	Reading			0.72	0.90	0.61	0.87
	Writing				0.71	0.55	0.86
	Comprehension					0.75	0.92
	Oral						0.85
5	Speaking	0.47	0.48	0.44	0.52	0.88	0.72
	Listening		0.62	0.53	0.85	0.79	0.79
	Reading			0.70	0.90	0.60	0.87
	Writing				0.69	0.53	0.85
	Comprehension					0.74	0.91
	Oral						0.84

Table 29. CELApro Scale Score Correlations, Grade Span 6–8

		Listening	Reading	Writing	Comprehension	Oral	Total
6	Speaking	0.47	0.45	0.50	0.52	0.90	0.75
	Listening		0.56	0.51	0.78	0.75	0.80
	Reading			0.67	0.92	0.56	0.82
	Writing				0.69	0.57	0.84
	Comprehension					0.70	0.90
	Oral						0.87
7	Speaking	0.52	0.47	0.51	0.54	0.91	0.77
	Listening		0.57	0.54	0.78	0.78	0.82
	Reading			0.67	0.92	0.56	0.82
	Writing				0.70	0.58	0.84
	Comprehension					0.71	0.90
	Oral						0.88
8	Speaking	0.54	0.53	0.55	0.59	0.91	0.80
	Listening		0.58	0.54	0.78	0.79	0.82
	Reading			0.67	0.93	0.60	0.84
	Writing				0.69	0.58	0.83
	Comprehension					0.73	0.90
	Oral						0.88

Table 30. CELApro Scale Score Correlations, Grade Span 9–12

		Listening	Reading	Writing	Comprehension	Oral	Total
9	Speaking	0.56	0.51	0.60	0.59	0.90	0.82
	Listening		0.64	0.61	0.83	0.83	0.84
	Reading			0.67	0.92	0.63	0.82
	Writing				0.72	0.67	0.86
	Comprehension					0.77	0.90
	Oral						0.92
10	Speaking	0.57	0.54	0.62	0.61	0.89	0.82
	Listening		0.66	0.64	0.84	0.84	0.86
	Reading			0.70	0.93	0.64	0.84
	Writing				0.74	0.68	0.87
	Comprehension					0.77	0.91
	Oral						0.91
11	Speaking	0.58	0.54	0.59	0.59	0.89	0.82
	Listening		0.68	0.62	0.85	0.83	0.86
	Reading			0.68	0.93	0.65	0.83
	Writing				0.70	0.65	0.85
	Comprehension					0.77	0.90
	Oral						0.91
12	Speaking	0.57	0.54	0.59	0.59	0.88	0.81
	Listening		0.70	0.64	0.85	0.84	0.87
	Reading			0.69	0.94	0.66	0.85
	Writing				0.72	0.65	0.85
	Comprehension					0.76	0.91
	Oral						0.91

Overall, the pattern of correlations among the four content domains of Listening, Speaking, Reading, and Writing is similar to the pattern observed in the 2009 data and is consistent with theoretical expectations for the CELApro language constructs. For example, the correlations support the distinction between the receptive language skills (Listening and Reading) and the productive language skills (Speaking and Writing). At all grade levels, the component exhibiting the highest correlation with the Listening scale is the Reading scale. The relationship between the productive skills of Speaking and Writing is less clear.

Consistent with last year, the highest single correlation coefficient among the four domains at each grade level is the correlation between the two orthographic domains of Reading and Writing.

Part 8. Special Studies

No special studies were conducted in the 2009–2010 administration.

References

American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Psychological Association.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing.* Oxford, England: Oxford University Press.

Bachman, L. E, and Palmer, A.S. (1996). *Language Testing in Practice*. Oxford, England: Oxford University Press.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37*, 29–51.

Bock, R. D., and Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika, 46,* 443–459.

Burket, G. R. (2002). *PARDUX* [Computer program]. Monterey, CA: CTB/McGraw-Hill.

Camilli, G., and Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. Thousand Oaks: Sage.

Chamot, A. U., and O'Malley, J. M. (1994). *The CALLA Handbook*. Reading, MA: Addison-Wesley.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

CTB/McGraw-Hill. (2007). *LAS Links Technical Manual*. Monterey, CA: Author.

Fleiss, J. L. and Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement, 33*, 613–619.

Green, D. R. (1975). 'What does it mean to say a test is biased?' Paper presented at American Educational Research Association, Washington, D.C.

Linn, R. L., and Harnisch, D. (1981). Interactions between item content and group membership in achievement test items. *Journal of Educational Measurement*, *18*, 109–118.

Linn, R. L., Levine, M. V., Hastings, C. N., and Wardrop, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement, 5*, 159–173.

Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems 71*, 179–181. Hillsdale, NJ: Lawrence Erlbaum.

Lord, F. M., and Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

MacMillan/McGraw-Hill (1993a). *Guidelines for Bias-Free Publishing.* New York: Author.

MacMillan/McGraw-Hill (1993b). Reflecting Diversity: Multicultural Guidelines for Educational Publishing Professionals. New York: Author.

O'Malley, J. M., and Valdez Pierce, L. (1996). *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. Boston: Addison-Wesley.

Sandoval, J. H., and Mille, M. P. (1979, August). "Accuracy of judgments of WISC-R item difficulty for minority groups." Paper presented at the annual meeting of the American Psychological Association, New York.

Savignon, S. (1972). *Communicative Competence: An Experiment in Foreign Language Teaching*. Philadelphia: The Center for Curriculum Development.

Savignon, S. (1997). *Communicative Competence: Theory and Classroom Practice* (2nd ed.). New York: McGraw-Hill.

Scheuneman, J.D. (1984). A theoretical framework for the exploration of causes and effects of bias in testing. *Educational Psychology, 19*(4), 219–225.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction*, 3–32. Cambridge, UK: Cambridge University Press.

Stocking, M.L. and Lord, F.M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*, 201–210.

Subkoviak, M. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement, 25(1)*, 47–55.

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix A: Item Analysis Results

List of Tables

Table 1A. Item Analysis Grade Span 1: K, Speaking

Item Number	Item Type	*P*-Value	Item-Total Correlation	Point Biserial at A	at B	at C	at D	Percent Omit
1	DCR	0.92	0.40	0.40	-0.25	-0.31		0.10%
2	DCR	0.62	0.58	0.58	-0.26	-0.54		0.17%
3	DCR	0.64	0.59	0.59	-0.19	-0.51		0.52%
4	DCR	0.69	0.57	0.57	-0.25	-0.51		2.59%
5	DCR	0.58	0.48	0.48	-0.14	-0.47		3.08%
6	DCR	0.44	0.55	0.56	-0.17	-0.54		3.53%
7	DCR	0.35	0.46	0.46	-0.17	-0.31		3.73%
8	DCR	0.33	0.54	0.54	-0.16	-0.44		3.99%
9	DCR	0.09	0.31	0.31	0.09	-0.31		3.99%
10	DCR	0.19	0.44	0.44	0.01	-0.44		4.00%
11	CR	0.67	0.80	-0.56	-0.21	0.61		0.15%
12	CR	0.67	0.83	-0.65	-0.12	0.63		0.45%
13	CR	0.69	0.74	-0.57	-0.18	0.60		6.06%
14	CR	0.68	0.77	-0.55	-0.24	0.61		6.31%
15	CR	0.50	0.76	-0.60	0.15	0.50		6.52%
16	CR	0.59	0.75	-0.56	0.03	0.52		0.44%
17	CR	0.54	0.74	-0.55	-0.05	0.56		0.56%
18	CR	0.44	0.67	-0.56	0.13	0.51		10.21%
19	CR	0.46	0.75	-0.61	0.07	0.57		10.32%
20	CR	0.54	0.81	-0.60	-0.20	0.32	0.49	0.62%

Table 2A. Item Analysis Grade Span 1: K, Listening

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.46	0.44	-0.38	0.45	-0.11		3.90%
2	MC	0.59	0.52	-0.43	-0.20	0.52		4.09%
3	MC	0.74	0.52	-0.47	0.52	-0.17		2.40%
4	MC	0.78	0.19	0.19	-0.15	-0.11		3.74%
5	MC	0.58	0.53	-0.47	-0.16	0.53		4.47%
6	MC	0.45	0.44	-0.31	-0.20	0.44		3.96%
7	MC	0.65	0.27	0.27	-0.25	-0.09		4.24%
8	MC	0.65	0.27	0.27	-0.24	-0.10		4.54%
9	MC	0.59	0.46	-0.43	-0.18	0.46		4.22%
10	MC	0.34	0.34	-0.27	0.34	-0.08		4.59%
11	MC	0.31	0.27	-0.36	0.27	0.09		5.29%
12	MC	0.39	0.38	-0.34	0.38	-0.05		5.38%
13	MC	0.36	0.40	-0.32	-0.09	0.40		7.63%
14	MC	0.31	0.37	-0.30	0.37	-0.06		5.90%
15	MC	0.40	0.39	-0.31	0.39	-0.10		7.56%
16*	MC	0.43	0.04	0.04	-0.10	0.05		8.41%
17	MC	0.23	0.26	-0.22	0.26	-0.01		8.49%
18	MC	0.57	0.27	0.27	-0.20	-0.14		4.06%
19	MC	0.49	0.44	-0.40	-0.10	0.44		5.22%
20	MC	0.39	0.39	-0.28	0.39	-0.14		6.01%
21	MC	0.63	0.41	-0.36	0.41	-0.15		6.70%

*This item was removed prior to scoring.

Table 3A. Item Analysis Grade Span 1: K, Reading

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial at A	at B	at C	at D	Percent Omit
1	MC	0.76	0.39	-0.35	0.39	-0.14		4.02%
2	MC	0.89	0.40	-0.34	0.40	-0.20		2.22%
3	MC	0.85	0.44	-0.35	-0.25	0.44		3.33%
4	MC	0.78	0.47	-0.39	-0.22	0.47		4.35%
5	MC	0.76	0.47	-0.37	0.47	-0.26		3.54%
6	MC	0.57	0.17	0.17	-0.20	-0.04		5.95%
7	MC	0.44	0.43	-0.27	-0.22	0.43		6.79%
8	MC	0.39	0.21	0.21	-0.16	-0.11		6.65%
9	MC	0.35	0.36	-0.32	-0.04	0.36		7.71%
10	MC	0.48	0.39	-0.41	0.39	-0.02		7.33%
11	MC	0.35	0.41	-0.31	-0.11	0.41		7.92%
12	MC	0.67	0.31	0.31	-0.24	-0.15		3.49%
13	MC	0.80	0.35	0.36	-0.27	-0.21		3.81%
14	MC	0.36	0.44	-0.28	-0.22	0.44		5.06%
15	MC	0.44	0.34	-0.28	0.34	-0.08		6.97%
16	MC	0.35	0.48	-0.36	-0.16	0.48		7.64%
17	MC	0.56	0.27	0.27	-0.22	-0.11		8.78%
18	MC	0.33	0.44	-0.35	-0.09	0.44		8.36%
19	MC	0.55	0.47	-0.42	0.47	-0.12		8.00%
20	MC	0.25	0.16	-0.21	0.16	0.08		18.55%
21	MC	0.35	-0.01	-0.01	-0.02	0.03		21.60%
22	MC	0.68	0.44	-0.41	0.44	-0.14		3.21%
23	MC	0.61	0.40	-0.30	-0.23	0.40		3.19%
24	MC	0.73	0.45	-0.35	0.45	-0.23		3.87%
25	MC	0.35	0.41	-0.33	0.41	-0.10		11.59%
26	MC	0.35	0.22	-0.20	0.22	-0.02		11.78%
27	MC	0.53	0.18	0.18	-0.14	-0.07		5.46%
28	MC	0.73	0.44	-0.39	0.44	-0.18		6.13%
29	MC	0.71	0.41	-0.33	-0.22	0.41		5.84%
30	MC	0.33	0.30	-0.25	0.30	-0.05		12.49%
31	MC	0.42	0.42	-0.29	-0.18	0.42		10.41%

Table 4A. Item Analysis Grade Span 1: K, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.19	0.23	-0.15	-0.04	0.23		2.58%
2	MC	0.41	0.19	0.19	-0.12	-0.09		5.66%
3	MC	0.20	0.23	-0.08	-0.11	0.23		7.36%
4	MC	0.32	0.14	-0.06	0.14	-0.08		10.23%
5	MC	0.27	0.18	-0.13	-0.05	0.18		14.04%
6	MC	0.49	0.23	0.23	-0.19	-0.08		15.27%
7	MC	0.42	0.24	-0.15	0.24	-0.14		15.98%
8	MC	0.39	0.13	0.13	-0.09	-0.05		15.80%
9	MC	0.47	0.12	0.12	-0.08	-0.06		17.86%
10	MC	0.40	0.21	0.21	-0.17	-0.07		18.09%
11	MC	0.23	0.19	0.00	0.19	-0.18		18.48%
12	MC	0.23	0.07	-0.05	0.07	-0.02		18.31%
13	MC	0.34	0.11	0.11	-0.17	0.05		18.83%
14	MC	0.15	0.14	0.05	-0.17	0.14		20.08%
15	MC	0.39	0.23	0.23	-0.19	-0.08		21.23%
16	MC	0.34	0.21	0.21	-0.10	-0.14		23.47%
17	MC	0.10	0.14	-0.04	-0.06	0.15		21.54%
18	MC	0.36	0.21	0.21	-0.16	-0.08		22.66%
19	MC	0.32	0.22	0.22	-0.11	-0.13		23.23%
20	MC	0.18	0.14	-0.02	-0.11	0.14		23.08%
21	CR	0.25	0.55	-0.50	0.37	0.33		7.21%
22	CR	0.24	0.56	-0.51	0.36	0.35		11.55%
23	CR	0.30	0.60	-0.55	0.27	0.45		4.85%
24	CR	0.28	0.61	-0.56	0.33	0.43		7.80%
25	CR	0.28	0.61	-0.55	0.30	0.45		8.45%

Table 5A. Item Analysis Grade Span 1: K, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.46	0.39	-0.33	0.39	-0.10		3.86%
2	MC	0.59	0.49	-0.40	-0.19	0.49		4.05%
3	MC	0.75	0.49	-0.44	0.49	-0.17		2.33%
4	MC	0.79	0.17	0.17	-0.13	-0.10		3.66%
5	MC	0.58	0.49	-0.43	-0.15	0.49		4.36%
6	MC	0.45	0.38	-0.25	-0.20	0.38		3.85%
7	MC	0.65	0.26	0.26	-0.24	-0.07		4.13%
8	MC	0.65	0.25	0.25	-0.23	-0.08		4.43%
9	MC	0.59	0.44	-0.41	-0.16	0.44		4.11%
10	MC	0.31	0.21	-0.32	0.21	0.11		5.16%
11	MC	0.39	0.33	-0.31	0.33	-0.04		5.24%
12	MC	0.36	0.36	-0.30	-0.07	0.36		7.47%
13	MC	0.31	0.32	-0.28	0.32	-0.04		5.75%
14	MC	0.40	0.34	-0.28	0.34	-0.09		7.39%
15	MC	0.43	0.00	0.00	-0.08	0.08		8.25%
16	MC	0.23	0.21	-0.19	0.21	0.01		8.32%
17	MC	0.57	0.23	0.23	-0.17	-0.13		3.90%
18	MC	0.49	0.40	-0.37	-0.08	0.40		5.05%
19	MC	0.40	0.34	-0.25	0.34	-0.11		5.84%
20	MC	0.63	0.37	-0.34	0.37	-0.13		6.54%
21	MC	0.77	0.39	-0.36	0.39	-0.13		3.97%
22	MC	0.89	0.38	-0.32	0.38	-0.19		2.17%
23	MC	0.85	0.40	-0.34	-0.21	0.40		3.27%
24	MC	0.78	0.42	-0.36	-0.19	0.42		4.28%
25	MC	0.39	0.17	0.17	-0.12	-0.09		6.54%
26	MC	0.35	0.31	-0.30	-0.02	0.32		7.56%
27	MC	0.48	0.38	-0.40	0.38	-0.02		7.20%
28	MC	0.35	0.35	-0.29	-0.07	0.35		7.79%
29	MC	0.25	0.16	-0.21	0.16	0.07		18.40%
30	MC	0.35	-0.03	-0.03	0.01	0.03		21.43%
31	MC	0.68	0.45	-0.42	0.45	-0.14		3.11%
32	MC	0.61	0.42	-0.33	-0.23	0.42		3.08%
33	MC	0.73	0.48	-0.38	0.48	-0.24		3.74%
34	MC	0.35	0.37	-0.31	0.37	-0.08		11.44%
35	MC	0.35	0.22	-0.21	0.22	-0.02		11.61%
36	MC	0.53	0.17	0.17	-0.14	-0.06		5.30%
37	MC	0.73	0.47	-0.42	0.47	-0.17		5.97%
38	MC	0.71	0.43	-0.35	-0.21	0.43		5.70%
39	MC	0.33	0.27	-0.24	0.27	-0.04		12.32%
40	MC	0.42	0.38	-0.28	-0.15	0.38		10.24%

Table 6A. Item Analysis Grade Span 1: K, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial at A	at B	at C	at D	Percent Omit
1	DCR	0.93	0.38	0.38	-0.24	-0.29		0.09%
2	DCR	0.62	0.56	0.56	-0.26	-0.51		0.17%
3	DCR	0.64	0.58	0.58	-0.20	-0.50		0.50%
4	DCR	0.69	0.54	0.54	-0.24	-0.48		2.51%
5	DCR	0.58	0.45	0.45	-0.14	-0.45		2.98%
6	DCR	0.44	0.54	0.54	-0.18	-0.51		3.37%
7	DCR	0.35	0.47	0.47	-0.20	-0.29		3.56%
8	DCR	0.33	0.54	0.54	-0.19	-0.42		3.79%
9	DCR	0.09	0.33	0.33	0.06	-0.29		3.79%
10	DCR	0.20	0.43	0.43	-0.02	-0.42		3.80%
11	CR	0.68	0.77	-0.53	-0.21	0.59		0.14%
12	CR	0.68	0.80	-0.61	-0.13	0.61		0.42%
13	CR	0.69	0.70	-0.54	-0.17	0.57		5.80%
14	CR	0.68	0.73	-0.52	-0.23	0.58		6.02%
15	CR	0.51	0.72	-0.56	0.13	0.48		6.23%
16	CR	0.59	0.71	-0.52	0.03	0.49		0.39%
17	CR	0.54	0.72	-0.53	-0.05	0.55		0.50%
18	CR	0.44	0.62	-0.51	0.11	0.48		9.88%
19	CR	0.46	0.73	-0.58	0.07	0.54		9.98%
20	CR	0.55	0.78	-0.57	-0.19	0.30	0.47	0.57%
21	MC	0.46	0.30	-0.25	0.30	-0.09		3.90%
22	MC	0.59	0.37	-0.28	-0.18	0.37		4.08%
23	MC	0.74	0.46	-0.40	0.46	-0.20		2.39%
24	MC	0.78	0.19	0.19	-0.15	-0.11		3.75%
25	MC	0.58	0.39	-0.33	-0.14	0.39		4.46%
26	MC	0.45	0.27	-0.15	-0.18	0.27		3.95%
27	MC	0.65	0.23	0.23	-0.20	-0.08		4.23%
28	MC	0.65	0.25	0.25	-0.22	-0.08		4.54%
29	MC	0.59	0.44	-0.34	-0.22	0.44		4.20%
30	MC	0.34	0.20	-0.14	0.20	-0.07		4.57%
31	MC	0.31	0.15	-0.23	0.15	0.08		5.28%
32	MC	0.39	0.25	-0.23	0.25	-0.02		5.39%
33	MC	0.36	0.29	-0.21	-0.09	0.29		7.60%
34	MC	0.31	0.25	-0.24	0.25	-0.01		5.87%
35	MC	0.40	0.25	-0.18	0.25	-0.09		7.55%
36	MC	0.43	0.01	0.01	-0.08	0.07		8.40%
37	MC	0.23	0.13	-0.12	0.13	0.01		8.46%
38	MC	0.57	0.21	0.21	-0.17	-0.10		4.04%
39	MC	0.49	0.36	-0.31	-0.09	0.36		5.16%
40	MC	0.39	0.26	-0.18	0.27	-0.11		5.97%
41	MC	0.63	0.32	-0.28	0.32	-0.12		6.64%

Table 7A. Item Analysis Grade Span 1: 1, Speaking

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.97	0.33	0.33	-0.22	-0.27		0.08%
2	DCR	0.75	0.50	0.50	-0.31	-0.47		0.28%
3	DCR	0.82	0.57	0.57	-0.23	-0.50		0.51%
4	DCR	0.85	0.51	0.52	-0.27	-0.47		0.77%
5	DCR	0.70	0.43	0.43	-0.18	-0.43		0.90%
6	DCR	0.64	0.51	0.52	-0.28	-0.49		0.99%
7	DCR	0.66	0.48	0.48	-0.22	-0.40		1.12%
8	DCR	0.65	0.54	0.54	-0.31	-0.45		1.20%
9	DCR	0.26	0.39	0.39	-0.12	-0.29		1.20%
10	DCR	0.39	0.47	0.47	-0.20	-0.41		1.18%
11	CR	0.84	0.76	-0.49	-0.39	0.61		0.31%
12	CR	0.83	0.77	-0.56	-0.33	0.60		0.50%
13	CR	0.86	0.67	-0.48	-0.36	0.57		1.82%
14	CR	0.84	0.72	-0.50	-0.35	0.58		1.86%
15	CR	0.69	0.73	-0.56	-0.09	0.52		1.93%
16	CR	0.73	0.68	-0.51	-0.10	0.50		0.60%
17	CR	0.76	0.71	-0.50	-0.24	0.57		0.46%
18	CR	0.59	0.63	-0.52	-0.05	0.53		3.39%
19	CR	0.67	0.75	-0.58	-0.15	0.59		3.44%
20	CR	0.72	0.79	-0.50	-0.39	0.05	0.54	0.60%

Table 8A. Item Analysis Grade Span 1: 1, Listening

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.76	0.48	-0.39	0.48	-0.23		2.55%
2	MC	0.86	0.46	-0.39	-0.23	0.46		2.70%
3	MC	0.92	0.43	-0.38	0.43	-0.20		1.22%
4	MC	0.91	0.30	0.31	-0.26	-0.15		0.91%
5	MC	0.88	0.43	-0.39	-0.21	0.43		1.23%
6	MC	0.74	0.45	-0.35	-0.25	0.45		1.02%
7	MC	0.90	0.39	0.39	-0.35	-0.16		0.94%
8	MC	0.87	0.41	0.41	-0.33	-0.22		1.07%
9	MC	0.82	0.44	-0.34	-0.31	0.44		0.80%
10	MC	0.56	0.40	-0.27	0.40	-0.24		0.75%
11	MC	0.49	0.34	-0.28	0.34	-0.15		0.99%
12	MC	0.70	0.45	-0.37	0.45	-0.21		1.79%
13	MC	0.72	0.50	-0.35	-0.30	0.50		2.82%
14	MC	0.55	0.42	-0.29	0.42	-0.22		2.04%
15	MC	0.67	0.37	-0.28	0.37	-0.20		1.48%
16	MC	0.49	0.29	0.29	-0.20	-0.15		1.58%
17	MC	0.35	0.34	-0.25	0.34	-0.09		1.48%
18	MC	0.76	0.40	0.40	-0.35	-0.17		0.95%
19	MC	0.74	0.46	-0.34	-0.26	0.46		1.52%
20	MC	0.60	0.40	-0.26	0.40	-0.25		1.63%
21	MC	0.84	0.37	-0.30	0.37	-0.21		1.76%

Table 9A. Item Analysis Grade Span 1: 1, Reading

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.92	0.26	-0.22	0.26	-0.13		2.08%
2	MC	0.98	0.22	-0.18	0.22	-0.12		0.60%
3	MC	0.97	0.25	-0.20	-0.16	0.25		1.49%
4	MC	0.95	0.25	-0.19	-0.17	0.25		1.52%
5	MC	0.96	0.29	-0.21	0.29	-0.20		0.81%
6	MC	0.47	0.32	0.32	-0.16	-0.22		1.96%
7	MC	0.19	0.20	-0.02	-0.14	0.20		2.72%
8	MC	0.57	0.43	0.43	-0.15	-0.37		2.70%
9	MC	0.64	0.40	-0.27	-0.23	0.40		1.93%
10	MC	0.74	0.41	-0.32	0.41	-0.23		1.92%
11	MC	0.69	0.48	-0.31	-0.31	0.48		1.63%
12	MC	0.90	0.32	0.32	-0.27	-0.17		1.05%
13	MC	0.96	0.24	0.24	-0.19	-0.15		0.92%
14	MC	0.80	0.47	-0.41	-0.20	0.47		1.25%
15	MC	0.74	0.42	-0.35	0.42	-0.20		1.75%
16	MC	0.83	0.48	-0.32	-0.33	0.48		1.04%
17	MC	0.83	0.40	0.40	-0.32	-0.22		1.96%
18	MC	0.81	0.48	-0.36	-0.29	0.48		1.37%
19	MC	0.91	0.36	-0.31	0.36	-0.16		1.34%
20	MC	0.34	0.28	-0.24	0.28	-0.05		3.84%
21	MC	0.39	0.22	0.22	-0.15	-0.08		6.14%
22	MC	0.85	0.32	-0.23	0.32	-0.22		1.11%
23	MC	0.86	0.32	-0.20	-0.25	0.33		1.15%
24	MC	0.91	0.29	-0.21	0.29	-0.20		1.41%
25	MC	0.81	0.44	-0.39	0.44	-0.17		1.88%
26	MC	0.56	0.39	-0.26	0.39	-0.21		2.32%
27	MC	0.57	0.43	0.43	-0.33	-0.20		10.53%
28	MC	0.35	0.47	-0.27	-0.23	0.47		7.59%
29	MC	0.46	0.49	-0.28	-0.29	0.49		8.46%
30	MC	0.44	0.34	0.34	-0.15	-0.24		8.73%
31	MC	0.42	0.27	-0.21	0.27	-0.10		10.10%
32	MC	0.43	0.25	0.25	-0.18	-0.11		10.21%
33	MC	0.33	0.33	-0.17	-0.18	0.33		10.77%
34	MC	0.35	0.30	0.30	-0.07	-0.25		11.39%
35	MC	0.39	0.26	-0.14	0.26	-0.16		11.43%
36	MC	0.39	0.37	-0.17	-0.24	0.37		11.37%

Table 10A. Item Analysis Grade Span 1: 1, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.54	0.52	-0.36	-0.26	0.52		0.99%
2	MC	0.56	0.43	0.43	-0.25	-0.28		1.39%
3	MC	0.54	0.55	-0.31	-0.35	0.55		1.69%
4	MC	0.49	0.34	-0.20	0.34	-0.21		1.84%
5	MC	0.48	0.36	-0.22	-0.21	0.37		1.87%
6	MC	0.75	0.42	0.42	-0.32	-0.24		2.35%
7	MC	0.72	0.31	-0.24	0.31	-0.16		2.25%
8	MC	0.52	0.35	0.35	-0.23	-0.18		2.07%
9	MC	0.53	0.24	0.24	-0.17	-0.12		2.48%
10	MC	0.54	0.38	0.38	-0.29	-0.16		2.59%
11	MC	0.53	0.50	-0.25	0.50	-0.34		2.74%
12	MC	0.43	0.14	-0.06	0.14	-0.11		2.71%
13	MC	0.42	0.29	0.29	-0.21	-0.11		2.92%
14	MC	0.40	0.35	-0.11	-0.30	0.35		4.12%
15	MC	0.52	0.33	0.33	-0.26	-0.14		4.44%
16	MC	0.42	0.31	0.31	-0.13	-0.23		5.92%
17	MC	0.27	0.34	-0.17	-0.14	0.34		4.96%
18	MC	0.41	0.27	0.27	-0.18	-0.12		5.50%
19	MC	0.38	0.35	0.35	-0.11	-0.27		6.00%
20	MC	0.42	0.41	-0.25	-0.21	0.41		5.62%
21	CR	0.56	0.71	-0.61	0.24	0.45		1.88%
22	CR	0.59	0.73	-0.58	0.06	0.54		2.22%
23	CR	0.64	0.69	-0.57	0.01	0.51		1.11%
24	CR	0.63	0.69	-0.57	0.02	0.52		1.53%
25	CR	0.65	0.69	-0.57	-0.03	0.53		1.67%

Table 11A. Item Analysis Grade Span 1: 1, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.76	0.41	-0.34	0.41	-0.20		2.54%
2	MC	0.86	0.39	-0.33	-0.19	0.39		2.69%
3	MC	0.92	0.36	-0.31	0.36	-0.17		1.20%
4	MC	0.91	0.24	0.25	-0.22	-0.11		0.88%
5	MC	0.88	0.36	-0.33	-0.17	0.36		1.20%
6	MC	0.74	0.38	-0.30	-0.22	0.38		1.00%
7	MC	0.90	0.33	0.33	-0.30	-0.13		0.91%
8	MC	0.87	0.35	0.35	-0.29	-0.18		1.04%
9	MC	0.82	0.39	-0.29	-0.29	0.39		0.78%
10	MC	0.49	0.28	-0.23	0.28	-0.13		0.96%
11	MC	0.70	0.42	-0.33	0.42	-0.21		1.76%
12	MC	0.72	0.45	-0.32	-0.27	0.45		2.78%
13	MC	0.55	0.36	-0.26	0.36	-0.18		2.01%
14	MC	0.67	0.32	-0.24	0.32	-0.18		1.42%
15	MC	0.49	0.25	0.25	-0.19	-0.11		1.53%
16	MC	0.35	0.29	-0.23	0.29	-0.06		1.42%
17	MC	0.76	0.36	0.36	-0.31	-0.15		0.92%
18	MC	0.74	0.42	-0.31	-0.24	0.42		1.48%
19	MC	0.60	0.34	-0.21	0.34	-0.22		1.59%
20	MC	0.84	0.31	-0.25	0.31	-0.17		1.72%
21	MC	0.92	0.27	-0.23	0.27	-0.14		2.07%
22	MC	0.98	0.21	-0.18	0.22	-0.12		0.57%
23	MC	0.97	0.24	-0.20	-0.14	0.24		1.47%
24	MC	0.95	0.24	-0.18	-0.16	0.24		1.49%
25	MC	0.57	0.42	0.42	-0.15	-0.36		2.66%
26	MC	0.64	0.37	-0.25	-0.21	0.37		1.88%
27	MC	0.74	0.40	-0.32	0.40	-0.22		1.87%
28	MC	0.69	0.44	-0.29	-0.28	0.44		1.58%
29	MC	0.34	0.25	-0.22	0.25	-0.04		3.78%
30	MC	0.39	0.19	0.19	-0.14	-0.06		6.08%
31	MC	0.85	0.34	-0.25	0.34	-0.22		1.06%
32	MC	0.86	0.34	-0.21	-0.26	0.34		1.11%
33	MC	0.91	0.31	-0.23	0.31	-0.21		1.37%
34	MC	0.81	0.41	-0.37	0.41	-0.16		1.83%
35	MC	0.56	0.37	-0.25	0.37	-0.20		2.27%
36	MC	0.57	0.40	0.40	-0.31	-0.19		10.48%
37	MC	0.35	0.43	-0.25	-0.20	0.43		7.52%
38	MC	0.46	0.46	-0.26	-0.27	0.46		8.39%
39	MC	0.44	0.32	0.32	-0.14	-0.23		8.66%
40	MC	0.42	0.25	-0.18	0.25	-0.10		10.04%
41	MC	0.43	0.22	0.22	-0.17	-0.10		10.15%
42	MC	0.33	0.31	-0.17	-0.16	0.31		10.71%
43	MC	0.35	0.27	0.27	-0.07	-0.22		11.33%
44	MC	0.39	0.23	-0.13	0.23	-0.13		11.37%
45	MC	0.39	0.33	-0.15	-0.22	0.33		11.31%

Table 12A. Item Analysis Grade Span 1: 1, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.97	0.31	0.31	-0.21	-0.25		0.08%
2	DCR	0.75	0.47	0.47	-0.30	-0.44		0.28%
3	DCR	0.82	0.56	0.56	-0.24	-0.49		0.51%
4	DCR	0.85	0.49	0.50	-0.27	-0.44		0.70%
5	DCR	0.70	0.40	0.40	-0.16	-0.41		0.82%
6	DCR	0.64	0.49	0.49	-0.27	-0.46		0.90%
7	DCR	0.66	0.49	0.49	-0.24	-0.39		1.02%
8	DCR	0.65	0.54	0.54	-0.32	-0.42		1.10%
9	DCR	0.26	0.41	0.41	-0.15	-0.27		1.08%
10	DCR	0.39	0.47	0.47	-0.21	-0.39		1.06%
11	CR	0.85	0.73	-0.47	-0.39	0.59		0.25%
12	CR	0.83	0.74	-0.53	-0.33	0.58		0.43%
13	CR	0.86	0.64	-0.45	-0.34	0.54		1.68%
14	CR	0.84	0.69	-0.47	-0.34	0.55		1.72%
15	CR	0.69	0.70	-0.54	-0.08	0.50		1.79%
16	CR	0.73	0.63	-0.46	-0.10	0.45		0.53%
17	CR	0.76	0.69	-0.48	-0.23	0.55		0.39%
18	CR	0.60	0.57	-0.47	-0.05	0.49		3.23%
19	CR	0.67	0.71	-0.54	-0.14	0.56		3.29%
20	CR	0.72	0.75	-0.46	-0.38	0.04	0.52	0.55%
21	MC	0.76	0.36	-0.28	0.36	-0.19		2.54%
22	MC	0.86	0.34	-0.28	-0.19	0.35		2.68%
23	MC	0.92	0.37	-0.31	0.37	-0.19		1.23%
24	MC	0.91	0.26	0.26	-0.22	-0.13		0.91%
25	MC	0.88	0.32	-0.29	-0.15	0.32		1.23%
26	MC	0.74	0.31	-0.24	-0.20	0.31		1.01%
27	MC	0.90	0.32	0.32	-0.29	-0.14		0.94%
28	MC	0.87	0.36	0.36	-0.30	-0.18		1.07%
29	MC	0.82	0.46	-0.30	-0.36	0.46		0.81%
30	MC	0.56	0.29	-0.16	0.29	-0.23		0.75%
31	MC	0.49	0.25	-0.20	0.25	-0.12		0.99%
32	MC	0.70	0.35	-0.28	0.35	-0.16		1.79%
33	MC	0.72	0.39	-0.26	-0.26	0.39		2.81%
34	MC	0.55	0.28	-0.23	0.28	-0.11		2.04%
35	MC	0.67	0.27	-0.19	0.27	-0.18		1.47%
36	MC	0.49	0.18	0.19	-0.16	-0.07		1.57%
37	MC	0.35	0.24	-0.18	0.24	-0.06		1.46%
38	MC	0.76	0.31	0.31	-0.27	-0.14		0.96%
39	MC	0.74	0.38	-0.29	-0.21	0.38		1.53%
40	MC	0.60	0.30	-0.18	0.30	-0.21		1.63%
41	MC	0.84	0.30	-0.26	0.30	-0.15		1.76%

Table 13A. Item Analysis Grade Span 1: 2, Speaking

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.98	0.35	0.35	-0.23	-0.30		0.05%
2	DCR	0.82	0.48	0.48	-0.34	-0.46		0.17%
3	DCR	0.90	0.55	0.55	-0.26	-0.48		0.38%
4	DCR	0.91	0.48	0.48	-0.29	-0.44		0.47%
5	DCR	0.76	0.46	0.46	-0.28	-0.38		0.63%
6	DCR	0.73	0.49	0.49	-0.34	-0.43		0.76%
7	DCR	0.83	0.43	0.43	-0.26	-0.33		0.71%
8	DCR	0.81	0.50	0.50	-0.36	-0.38		0.72%
9	DCR	0.44	0.41	0.41	-0.24	-0.26		0.78%
10	DCR	0.54	0.49	0.49	-0.34	-0.34		0.84%
11	CR	0.92	0.73	-0.47	-0.42	0.59		0.18%
12	CR	0.90	0.72	-0.51	-0.38	0.57		0.45%
13	CR	0.93	0.62	-0.44	-0.38	0.53		0.89%
14	CR	0.91	0.66	-0.45	-0.37	0.53		0.92%
15	CR	0.77	0.69	-0.51	-0.23	0.53		1.01%
16	CR	0.79	0.64	-0.46	-0.20	0.51		0.38%
17	CR	0.87	0.69	-0.46	-0.31	0.55		0.28%
18	CR	0.70	0.65	-0.52	-0.17	0.56		1.55%
19	CR	0.79	0.73	-0.53	-0.28	0.59		1.49%
20	CR	0.81	0.76	-0.46	-0.42	-0.13	0.56	0.44%

Table 14A. Item Analysis Grade Span 1: 2, Listening

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.93	0.46	-0.36	0.46	-0.29		0.96%
2	MC	0.95	0.37	-0.29	-0.22	0.37		1.12%
3	MC	0.97	0.39	-0.34	0.39	-0.19		0.43%
4	MC	0.97	0.31	0.31	-0.24	-0.19		0.38%
5	MC	0.96	0.37	-0.32	-0.23	0.37		0.40%
6	MC	0.90	0.40	-0.32	-0.24	0.40		0.39%
7	MC	0.97	0.33	0.33	-0.28	-0.17		0.28%
8	MC	0.96	0.38	0.38	-0.31	-0.23		0.33%
9	MC	0.91	0.41	-0.31	-0.33	0.41		0.27%
10	MC	0.72	0.38	-0.26	0.38	-0.30		0.34%
11	MC	0.64	0.37	-0.20	0.37	-0.28		0.38%
12	MC	0.91	0.42	-0.33	0.42	-0.25		0.46%
13	MC	0.90	0.42	-0.30	-0.27	0.42		0.68%
14	MC	0.72	0.42	-0.29	0.42	-0.26		0.57%
15	MC	0.82	0.34	-0.27	0.34	-0.19		0.48%
16	MC	0.62	0.38	0.38	-0.18	-0.29		0.59%
17	MC	0.48	0.38	-0.28	0.38	-0.16		0.54%
18	MC	0.87	0.41	0.41	-0.38	-0.16		0.37%
19	MC	0.87	0.45	-0.35	-0.26	0.45		0.49%
20	MC	0.73	0.41	-0.27	0.41	-0.28		0.64%
21	MC	0.93	0.34	-0.30	0.34	-0.18		0.58%

Table 15A. Item Analysis Grade Span 1: 2, Reading

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.96	0.25	-0.17	0.25	-0.19		0.97%
2	MC	0.99	0.16	-0.13	0.16	-0.10		0.24%
3	MC	0.99	0.18	-0.13	-0.13	0.18		0.50%
4	MC	0.98	0.21	-0.16	-0.15	0.21		0.40%
5	MC	0.99	0.24	-0.17	0.25	-0.18		0.23%
6	MC	0.58	0.45	0.45	-0.18	-0.37		0.66%
7	MC	0.32	0.37	-0.11	-0.27	0.37		0.89%
8	MC	0.73	0.46	0.46	-0.15	-0.42		0.85%
9	MC	0.78	0.42	-0.29	-0.27	0.42		0.47%
10	MC	0.88	0.42	-0.30	0.42	-0.29		0.40%
11	MC	0.85	0.48	-0.33	-0.32	0.48		0.38%
12	MC	0.95	0.28	0.29	-0.23	-0.17		0.39%
13	MC	0.99	0.21	0.21	-0.17	-0.12		0.38%
14	MC	0.94	0.39	-0.34	-0.18	0.39		0.53%
15	MC	0.90	0.45	-0.38	0.45	-0.22		0.46%
16	MC	0.95	0.42	-0.29	-0.30	0.42		0.37%
17	MC	0.95	0.39	0.39	-0.29	-0.25		0.77%
18	MC	0.95	0.41	-0.31	-0.26	0.41		0.34%
19	MC	0.98	0.28	-0.25	0.28	-0.14		0.31%
20	MC	0.63	0.52	-0.31	0.52	-0.35		0.89%
21	MC	0.58	0.42	0.42	-0.32	-0.18		1.49%
22	MC	0.91	0.34	-0.22	0.34	-0.25		0.74%
23	MC	0.95	0.28	-0.16	-0.23	0.28		0.69%
24	MC	0.95	0.26	-0.18	0.26	-0.20		0.80%
25	MC	0.95	0.34	-0.29	0.34	-0.18		1.00%
26	MC	0.77	0.49	-0.34	0.49	-0.31		0.93%
27	MC	0.82	0.48	0.49	-0.38	-0.28		6.52%
28	MC	0.67	0.55	-0.39	-0.31	0.55		1.71%
29	MC	0.79	0.57	-0.35	-0.41	0.57		1.95%
30	MC	0.67	0.45	0.45	-0.24	-0.33		2.06%
31	MC	0.64	0.45	-0.31	0.45	-0.26		3.51%
32	MC	0.70	0.53	0.53	-0.44	-0.23		2.49%
33	MC	0.61	0.51	-0.30	-0.34	0.51		2.90%
34	MC	0.59	0.49	0.49	-0.26	-0.36		3.19%
35	MC	0.54	0.31	-0.18	0.31	-0.20		3.06%
36	MC	0.68	0.53	-0.34	-0.35	0.53		3.09%

Table 16A. Item Analysis Grade Span 1: 2, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.82	0.49	-0.36	-0.32	0.49		0.43%
2	MC	0.79	0.43	0.43	-0.27	-0.35		0.70%
3	MC	0.84	0.59	-0.40	-0.40	0.59		0.55%
4	MC	0.70	0.49	-0.38	0.49	-0.24		0.69%
5	MC	0.72	0.47	-0.24	-0.37	0.47		0.63%
6	MC	0.91	0.43	0.43	-0.31	-0.29		0.89%
7	MC	0.84	0.36	-0.29	0.36	-0.20		0.85%
8	MC	0.72	0.41	0.42	-0.29	-0.25		0.74%
9	MC	0.70	0.37	0.37	-0.27	-0.24		0.47%
10	MC	0.70	0.42	0.42	-0.30	-0.24		0.58%
11	MC	0.78	0.54	-0.35	0.54	-0.38		0.57%
12	MC	0.51	0.17	-0.05	0.17	-0.19		0.55%
13	MC	0.58	0.37	0.37	-0.18	-0.27		0.95%
14	MC	0.66	0.44	-0.28	-0.33	0.44		1.68%
15	MC	0.69	0.41	0.41	-0.30	-0.23		1.16%
16	MC	0.62	0.45	0.45	-0.32	-0.25		1.55%
17	MC	0.50	0.45	-0.22	-0.30	0.45		1.00%
18	MC	0.57	0.35	0.35	-0.23	-0.21		1.20%
19	MC	0.59	0.49	0.49	-0.21	-0.38		1.42%
20	MC	0.69	0.52	-0.35	-0.32	0.52		1.26%
21	CR	0.68	0.74	-0.56	-0.08	0.55		0.64%
22	CR	0.66	0.69	-0.54	0.04	0.44		1.32%
23	CR	0.68	0.71	-0.58	0.01	0.48		0.54%
24	CR	0.66	0.69	-0.57	0.05	0.48		1.11%
25	CR	0.61	0.78	-0.45	-0.42	0.39	0.36	1.43%

Table 17A. Item Analysis Grade Span 1: 2, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.93	0.40	-0.30	0.40	-0.25		0.96%
2	MC	0.95	0.31	-0.25	-0.18	0.31		1.11%
3	MC	0.98	0.31	-0.27	0.31	-0.16		0.42%
4	MC	0.97	0.25	0.25	-0.20	-0.14		0.36%
5	MC	0.96	0.29	-0.26	-0.18	0.29		0.39%
6	MC	0.90	0.33	-0.27	-0.20	0.33		0.38%
7	MC	0.97	0.27	0.27	-0.24	-0.13		0.27%
8	MC	0.96	0.31	0.31	-0.26	-0.18		0.31%
9	MC	0.91	0.36	-0.24	-0.30	0.36		0.24%
10	MC	0.64	0.30	-0.15	0.30	-0.23		0.37%
11	MC	0.91	0.39	-0.31	0.39	-0.23		0.44%
12	MC	0.90	0.37	-0.25	-0.25	0.37		0.67%
13	MC	0.72	0.34	-0.24	0.34	-0.20		0.54%
14	MC	0.82	0.27	-0.21	0.27	-0.16		0.46%
15	MC	0.62	0.32	0.32	-0.16	-0.23		0.57%
16	MC	0.48	0.29	-0.24	0.29	-0.10		0.52%
17	MC	0.87	0.34	0.34	-0.31	-0.14		0.34%
18	MC	0.87	0.39	-0.30	-0.24	0.39		0.47%
19	MC	0.73	0.30	-0.19	0.30	-0.24		0.62%
20	MC	0.93	0.27	-0.25	0.27	-0.14		0.56%
21	MC	0.96	0.26	-0.18	0.26	-0.19		0.96%
22	MC	0.99	0.16	-0.13	0.17	-0.10		0.23%
23	MC	0.99	0.19	-0.14	-0.13	0.19		0.49%
24	MC	0.98	0.21	-0.16	-0.15	0.21		0.40%
25	MC	0.73	0.44	0.44	-0.15	-0.40		0.85%
26	MC	0.78	0.39	-0.27	-0.25	0.39		0.47%
27	MC	0.88	0.41	-0.31	0.41	-0.28		0.40%
28	MC	0.85	0.45	-0.31	-0.30	0.45		0.38%
29	MC	0.63	0.49	-0.30	0.49	-0.33		0.88%
30	MC	0.58	0.39	0.39	-0.30	-0.17		1.48%
31	MC	0.91	0.35	-0.23	0.35	-0.26		0.73%
32	MC	0.95	0.29	-0.17	-0.23	0.29		0.68%
33	MC	0.95	0.28	-0.19	0.28	-0.22		0.79%
34	MC	0.95	0.33	-0.28	0.33	-0.17		0.98%
35	MC	0.77	0.48	-0.34	0.48	-0.31		0.91%
36	MC	0.82	0.47	0.47	-0.37	-0.26		6.51%
37	MC	0.67	0.52	-0.36	-0.29	0.52		1.69%
38	MC	0.79	0.54	-0.34	-0.39	0.55		1.93%
39	MC	0.67	0.42	0.42	-0.22	-0.31		2.04%
40	MC	0.64	0.43	-0.28	0.43	-0.26		3.49%
41	MC	0.70	0.50	0.50	-0.41	-0.23		2.48%
42	MC	0.61	0.49	-0.29	-0.31	0.49		2.89%
43	MC	0.59	0.45	0.45	-0.24	-0.33		3.17%
44	MC	0.54	0.28	-0.17	0.28	-0.18		3.05%
45	MC	0.68	0.50	-0.32	-0.32	0.50		3.08%

Table 18A. Item Analysis Grade Span 1: 2, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.98	0.34	0.34	-0.23	-0.28		0.05%
2	DCR	0.82	0.45	0.45	-0.32	-0.44		0.17%
3	DCR	0.90	0.54	0.54	-0.27	-0.46		0.38%
4	DCR	0.91	0.46	0.47	-0.28	-0.41		0.45%
5	DCR	0.76	0.44	0.44	-0.27	-0.35		0.60%
6	DCR	0.73	0.47	0.47	-0.33	-0.41		0.74%
7	DCR	0.83	0.43	0.43	-0.27	-0.32		0.68%
8	DCR	0.81	0.49	0.49	-0.36	-0.36		0.69%
9	DCR	0.44	0.41	0.41	-0.25	-0.26		0.75%
10	DCR	0.54	0.48	0.48	-0.33	-0.33		0.80%
11	CR	0.92	0.70	-0.46	-0.40	0.57		0.18%
12	CR	0.90	0.69	-0.50	-0.36	0.54		0.45%
13	CR	0.93	0.60	-0.42	-0.36	0.51		0.84%
14	CR	0.91	0.64	-0.44	-0.35	0.51		0.87%
15	CR	0.77	0.65	-0.49	-0.21	0.49		0.96%
16	CR	0.79	0.60	-0.43	-0.18	0.47		0.36%
17	CR	0.87	0.68	-0.45	-0.30	0.54		0.27%
18	CR	0.70	0.61	-0.50	-0.15	0.53		1.48%
19	CR	0.79	0.71	-0.51	-0.26	0.56		1.43%
20	CR	0.81	0.73	-0.45	-0.40	-0.12	0.54	0.44%
21	MC	0.93	0.37	-0.27	0.37	-0.24		0.96%
22	MC	0.95	0.29	-0.21	-0.21	0.29		1.12%
23	MC	0.98	0.34	-0.29	0.34	-0.18		0.43%
24	MC	0.97	0.25	0.26	-0.21	-0.15		0.38%
25	MC	0.96	0.27	-0.24	-0.16	0.27		0.40%
26	MC	0.90	0.30	-0.23	-0.21	0.30		0.39%
27	MC	0.97	0.29	0.29	-0.24	-0.16		0.29%
28	MC	0.96	0.33	0.33	-0.28	-0.18		0.33%
29	MC	0.91	0.42	-0.30	-0.34	0.42		0.27%
30	MC	0.72	0.27	-0.16	0.27	-0.26		0.34%
31	MC	0.64	0.28	-0.15	0.28	-0.22		0.39%
32	MC	0.91	0.32	-0.25	0.32	-0.19		0.46%
33	MC	0.90	0.32	-0.21	-0.24	0.32		0.69%
34	MC	0.72	0.27	-0.22	0.27	-0.14		0.57%
35	MC	0.82	0.26	-0.20	0.26	-0.17		0.48%
36	MC	0.62	0.25	0.25	-0.13	-0.18		0.59%
37	MC	0.48	0.25	-0.20	0.25	-0.09		0.54%
38	MC	0.87	0.31	0.31	-0.27	-0.14		0.37%
39	MC	0.87	0.37	-0.29	-0.21	0.37		0.49%
40	MC	0.73	0.26	-0.15	0.26	-0.23		0.65%
41	MC	0.93	0.27	-0.28	0.27	-0.12		0.58%

Table 19A. Item Analysis Grade Span 2: 3–5, Speaking

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.98	0.25	0.25	-0.16	-0.27		0.06%
2	DCR	0.89	0.49	0.49	-0.36	-0.42		0.14%
3	DCR	0.76	0.51	0.51	-0.36	-0.35		0.15%
4	DCR	0.72	0.52	0.52	-0.40	-0.38		0.38%
5	DCR	0.95	0.35	0.35	-0.19	-0.29		0.44%
6	DCR	0.91	0.47	0.47	-0.35	-0.36		0.42%
7	DCR	0.87	0.49	0.49	-0.31	-0.38		0.44%
8	DCR	0.91	0.39	0.39	-0.29	-0.36		0.50%
9	DCR	0.54	0.44	0.44	-0.28	-0.27		0.53%
10	DCR	0.76	0.51	0.51	-0.38	-0.36		0.45%
11	CR	0.87	0.64	-0.44	-0.35	0.51		0.09%
12	CR	0.93	0.61	-0.42	-0.34	0.49		0.10%
13	CR	0.93	0.60	-0.41	-0.38	0.51		0.51%
14	CR	0.69	0.69	-0.55	-0.08	0.54		0.69%
15	CR	0.69	0.70	-0.55	-0.06	0.51		0.62%
16	CR	0.86	0.67	-0.45	-0.29	0.52		0.19%
17	CR	0.93	0.68	-0.39	-0.33	0.51		0.17%
18	CR	0.79	0.71	-0.51	-0.30	0.59		1.48%
19	CR	0.54	0.64	-0.54	0.07	0.49		1.57%
20	CR	0.83	0.74	-0.37	-0.44	-0.22	0.57	0.37%

Table 20A. Item Analysis Grade Span 2: 3–5, Listening

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.96	0.34	-0.27	0.35	-0.21		0.13%
2	MC	0.98	0.31	-0.21	-0.23	0.31		0.19%
3	MC	0.93	0.30	0.30	-0.29	-0.09		0.11%
4	MC	0.38	0.26	-0.07	-0.28	0.26		0.24%
5	MC	0.52	0.32	-0.26	0.32	-0.13		0.18%
6	MC	0.68	0.45	0.45	-0.32	-0.25		0.19%
7	MC	0.87	0.43	-0.34	-0.23	0.43		0.19%
8	MC	0.90	0.42	0.42	-0.27	-0.31		0.21%
9	MC	0.45	0.31	-0.16	0.31	-0.19		0.26%
10	MC	0.86	0.43	-0.30	-0.29	0.43		0.20%
11	MC	0.61	0.41	-0.36	0.41	-0.18		0.19%
12	MC	0.92	0.38	-0.22	0.38	-0.30		0.19%
13	MC	0.61	0.34	-0.06	-0.35	0.34		0.21%
14	MC	0.93	0.38	0.38	-0.30	-0.24		0.20%
15	MC	0.74	0.38	-0.35	-0.13	0.38		0.26%
16	MC	0.76	0.37	0.37	-0.22	-0.30		0.24%
17	MC	0.53	0.27	-0.12	-0.20	0.27		0.25%
18	MC	0.62	0.39	0.39	-0.31	-0.17		0.26%
19	MC	0.49	0.48	-0.37	0.48	-0.17		0.25%
20	MC	0.73	0.42	-0.26	0.42	-0.29		0.25%
21	MC	0.44	0.32	0.32	-0.13	-0.23		0.32%
22	MC	0.75	0.38	-0.26	0.38	-0.26		0.26%
23	MC	0.61	0.46	0.46	-0.33	-0.23		0.32%

Table 21A. Item Analysis Grade Span 2: 3–5, Reading

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.73	0.33	-0.28	0.33	-0.15		0.84%
2	MC	0.73	0.55	0.55	-0.44	-0.29		0.64%
3	MC	0.79	0.48	0.48	-0.26	-0.39		0.65%
4	MC	0.47	0.38	0.39	-0.09	-0.34		0.27%
5	MC	0.63	0.41	-0.23	-0.30	0.42		0.39%
6	MC	0.57	0.33	-0.25	-0.16	0.33		0.51%
7	MC	0.78	0.50	-0.38	-0.28	0.50		1.00%
8	MC	0.86	0.45	-0.30	-0.32	0.45		0.41%
9	MC	0.42	0.39	0.39	-0.30	-0.13		0.58%
10	MC	0.45	0.34	-0.18	-0.20	0.34		0.51%
11	MC	0.69	0.42	0.42	-0.35	-0.22		14.63%
12	MC	0.66	0.52	-0.29	-0.40	0.52		15.42%
13	MC	0.72	0.52	-0.44	0.52	-0.22		0.42%
14	MC	0.50	0.39	0.39	-0.24	-0.21		0.61%
15	MC	0.78	0.52	-0.35	0.52	-0.36		0.73%
16	MC	0.50	0.40	-0.22	0.40	-0.31		1.24%
17	MC	0.80	0.50	-0.33	-0.36	0.50		0.61%
18	MC	0.76	0.51	-0.33	-0.35	0.51		0.87%
19	MC	0.75	0.45	0.45	-0.29	-0.31		0.81%
20	MC	0.69	0.32	-0.18	0.32	-0.27		0.71%
21	MC	0.86	0.43	-0.36	0.43	-0.23		0.54%
22	MC	0.56	0.41	-0.29	0.41	-0.23		0.80%
23	MC	0.79	0.52	0.52	-0.33	-0.36		0.97%
24	MC	0.86	0.51	-0.39	0.51	-0.31		0.75%
25	MC	0.71	0.51	-0.31	0.51	-0.38		0.87%
26	MC	0.56	0.38	-0.24	-0.22	0.38	-0.09	0.67%
27	MC	0.67	0.44	-0.20	0.44	-0.17	-0.31	0.91%
28	MC	0.42	0.37	-0.30	-0.13	-0.14	0.37	1.49%
29	MC	0.77	0.51	-0.27	0.51	-0.22	-0.32	0.96%
30	MC	0.51	0.38	0.38	-0.18	-0.19	-0.16	1.16%
31	MC	0.51	0.45	-0.24	-0.26	0.45	-0.13	0.94%
32	MC	0.74	0.56	0.56	-0.34	-0.27	-0.26	1.12%
33	MC	0.68	0.53	-0.29	-0.25	-0.27	0.53	1.23%
34	MC	0.60	0.45	-0.24	0.45	-0.21	-0.20	0.97%
35	MC	0.83	0.54	-0.30	0.54	-0.31	-0.28	0.99%

Table 22A. Item Analysis Grade Span 2: 3–5, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.85	0.51	-0.43	0.51	-0.27		0.45%
2	MC	0.88	0.50	-0.27	-0.41	0.50		0.56%
3	MC	0.64	0.37	-0.30	0.37	-0.23		0.23%
4	MC	0.69	0.41	-0.33	-0.19	0.41		0.44%
5	MC	0.93	0.45	-0.34	-0.29	0.46		0.68%
6	MC	0.89	0.47	-0.30	-0.35	0.47		1.20%
7	MC	0.71	0.51	0.51	-0.38	-0.28		2.79%
8	MC	0.63	0.41	-0.27	-0.25	0.41		8.27%
9	MC	0.61	0.33	-0.14	0.33	-0.29		0.84%
10	MC	0.74	0.48	0.48	-0.35	-0.28		0.75%
11	MC	0.68	0.41	-0.24	0.41	-0.28		0.90%
12	MC	0.58	0.35	-0.29	-0.20	0.35		0.88%
13	MC	0.72	0.40	-0.25	0.40	-0.27		1.24%
14	MC	0.88	0.55	-0.31	-0.44	0.55		1.65%
15	MC	0.79	0.51	-0.35	0.51	-0.36		1.94%
16	MC	0.61	0.28	-0.07	-0.29	0.28		1.84%
17	MC	0.88	0.47	-0.31	0.47	-0.34		0.63%
18	MC	0.69	0.54	0.54	-0.30	-0.39		2.93%
19	MC	0.86	0.56	-0.38	-0.40	0.56		0.69%
20	MC	0.76	0.57	0.57	-0.38	-0.38		0.90%
21	MC	0.84	0.57	-0.40	0.57	-0.38		0.74%
22	CR	0.78	0.61	-0.40	-0.21	0.44		0.30%
23	CR	0.81	0.60	-0.40	-0.25	0.44		0.55%
24	CR	0.79	0.63	-0.45	-0.18	0.42		0.66%
25	CR	0.77	0.62	-0.44	-0.19	0.43		1.12%
26	CR	0.64	0.67	-0.30	-0.54	0.41	0.28	1.81%

Table 23A. Item Analysis Grade Span 2: 3–5, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.96	0.31	-0.25	0.31	-0.19		0.12%
2	MC	0.98	0.27	-0.19	-0.20	0.27		0.17%
3	MC	0.93	0.26	0.26	-0.25	-0.08		0.09%
4	MC	0.38	0.20	-0.02	-0.28	0.20		0.20%
5	MC	0.52	0.25	-0.24	0.25	-0.08		0.15%
6	MC	0.68	0.41	0.41	-0.29	-0.22		0.16%
7	MC	0.87	0.40	-0.33	-0.21	0.40		0.15%
8	MC	0.90	0.41	0.41	-0.25	-0.31		0.17%
9	MC	0.45	0.27	-0.14	0.27	-0.16		0.23%
10	MC	0.86	0.41	-0.29	-0.27	0.41		0.16%
11	MC	0.61	0.38	-0.33	0.38	-0.17		0.16%
12	MC	0.92	0.34	-0.20	0.34	-0.27		0.15%
13	MC	0.61	0.26	0.01	-0.31	0.26		0.18%
14	MC	0.93	0.35	0.35	-0.27	-0.21		0.16%
15	MC	0.74	0.32	-0.32	-0.08	0.32		0.22%
16	MC	0.76	0.32	0.32	-0.19	-0.28		0.21%
17	MC	0.53	0.19	-0.08	-0.14	0.19		0.20%
18	MC	0.62	0.34	0.34	-0.28	-0.13		0.21%
19	MC	0.49	0.43	-0.35	0.43	-0.14		0.21%
20	MC	0.44	0.25	0.25	-0.10	-0.18		0.27%
21	MC	0.76	0.32	-0.20	0.32	-0.25		0.22%
22	MC	0.61	0.38	0.38	-0.28	-0.18		0.28%
23	MC	0.46	0.34	-0.18	-0.19	0.34		0.51%
24	MC	0.69	0.41	0.41	-0.34	-0.21		14.59%
25	MC	0.66	0.50	-0.29	-0.38	0.50		15.38%
26	MC	0.72	0.50	-0.42	0.50	-0.22		0.41%
27	MC	0.50	0.39	0.39	-0.25	-0.21		0.61%
28	MC	0.78	0.50	-0.33	0.50	-0.35		0.73%
29	MC	0.50	0.40	-0.22	0.40	-0.30		1.23%
30	MC	0.80	0.47	-0.32	-0.33	0.47		0.60%
31	MC	0.76	0.48	-0.31	-0.32	0.48		0.87%
32	MC	0.75	0.44	0.44	-0.27	-0.31		0.80%
33	MC	0.69	0.30	-0.17	0.30	-0.26		0.70%
34	MC	0.86	0.42	-0.35	0.43	-0.22		0.52%
35	MC	0.56	0.42	-0.28	0.42	-0.24		0.78%
36	MC	0.79	0.51	0.51	-0.33	-0.36		0.95%
37	MC	0.86	0.50	-0.38	0.50	-0.30		0.74%
38	MC	0.71	0.49	-0.30	0.50	-0.37		0.85%
39	MC	0.56	0.38	-0.23	-0.21	0.38	-0.10	0.65%
40	MC	0.67	0.43	-0.20	0.43	-0.17	-0.30	0.89%
41	MC	0.43	0.36	-0.29	-0.13	-0.13	0.36	1.48%
42	MC	0.77	0.49	-0.26	0.49	-0.21	-0.31	0.94%
43	MC	0.51	0.38	0.38	-0.18	-0.18	-0.16	1.14%
44	MC	0.51	0.44	-0.23	-0.26	0.44	-0.12	0.93%
45	MC	0.75	0.54	0.54	-0.33	-0.26	-0.26	1.10%

Table 23A. Item Analysis Grade Span 2: 3–5, Comprehension (continued)

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
46	MC	0.68	0.50	-0.27	-0.23	-0.26	0.50	1.21%
47	MC	0.60	0.43	-0.23	0.43	-0.20	-0.20	0.95%
48	MC	0.83	0.52	-0.29	0.52	-0.30	-0.26	0.98%

Table 24A. Item Analysis Grade Span 2: 3–5, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.98	0.22	0.22	-0.14	-0.23		0.06%
2	DCR	0.89	0.45	0.45	-0.34	-0.38		0.13%
3	DCR	0.76	0.48	0.48	-0.34	-0.35		0.15%
4	DCR	0.73	0.47	0.47	-0.36	-0.35		0.34%
5	DCR	0.95	0.35	0.35	-0.20	-0.29		0.39%
6	DCR	0.91	0.44	0.44	-0.33	-0.33		0.37%
7	DCR	0.87	0.47	0.47	-0.30	-0.36		0.39%
8	DCR	0.91	0.36	0.36	-0.26	-0.33		0.45%
9	DCR	0.54	0.45	0.45	-0.30	-0.27		0.47%
10	DCR	0.76	0.48	0.48	-0.36	-0.33		0.40%
11	CR	0.87	0.58	-0.40	-0.31	0.46		0.08%
12	CR	0.93	0.56	-0.38	-0.32	0.46		0.09%
13	CR	0.93	0.54	-0.38	-0.34	0.45		0.44%
14	CR	0.69	0.66	-0.53	-0.06	0.51		0.62%
15	CR	0.69	0.67	-0.52	-0.05	0.48		0.55%
16	CR	0.86	0.63	-0.42	-0.27	0.49		0.17%
17	CR	0.93	0.63	-0.37	-0.31	0.47		0.16%
18	CR	0.79	0.67	-0.48	-0.27	0.55		1.40%
19	CR	0.54	0.61	-0.51	0.06	0.46		1.50%
20	CR	0.83	0.69	-0.33	-0.41	-0.21	0.54	0.37%
21	MC	0.96	0.39	-0.31	0.39	-0.23		0.13%
22	MC	0.98	0.34	-0.25	-0.24	0.34		0.19%
23	MC	0.93	0.31	0.31	-0.29	-0.13		0.11%
24	MC	0.38	0.17	0.00	-0.26	0.17		0.24%
25	MC	0.52	0.22	-0.20	0.22	-0.07		0.18%
26	MC	0.68	0.37	0.37	-0.26	-0.21		0.19%
27	MC	0.87	0.39	-0.30	-0.23	0.39		0.19%
28	MC	0.90	0.38	0.38	-0.24	-0.29		0.20%
29	MC	0.45	0.20	-0.10	0.20	-0.12		0.26%
30	MC	0.86	0.38	-0.26	-0.27	0.39		0.20%
31	MC	0.61	0.31	-0.26	0.32	-0.17		0.19%
32	MC	0.92	0.35	-0.20	0.35	-0.28		0.18%
33	MC	0.61	0.26	-0.01	-0.29	0.26		0.21%
34	MC	0.93	0.38	0.38	-0.29	-0.24		0.20%
35	MC	0.74	0.29	-0.27	-0.10	0.29		0.26%
36	MC	0.76	0.30	0.30	-0.16	-0.29		0.24%
37	MC	0.53	0.17	-0.08	-0.13	0.17		0.25%
38	MC	0.62	0.29	0.30	-0.23	-0.13		0.25%
39	MC	0.49	0.38	-0.30	0.38	-0.12		0.25%
40	MC	0.73	0.33	-0.22	0.33	-0.22		0.25%
41	MC	0.44	0.21	0.21	-0.08	-0.15		0.31%
42	MC	0.76	0.30	-0.18	0.30	-0.23		0.26%
43	MC	0.61	0.35	0.35	-0.26	-0.17		0.32%

Table 25A. Item Analysis Grade Span 3: 6–8, Speaking

Item Number	Item Type	*P*-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.82	0.57	0.57	-0.35	-0.44		0.11%
2	DCR	0.91	0.49	0.49	-0.33	-0.46		0.11%
3	DCR	0.31	0.34	0.34	-0.07	-0.28		0.30%
4	DCR	0.89	0.41	0.41	-0.28	-0.36		0.71%
5	DCR	0.74	0.46	0.46	-0.39	-0.25		0.88%
6	DCR	0.85	0.52	0.52	-0.43	-0.37		0.85%
7	DCR	0.76	0.49	0.49	-0.29	-0.39		0.92%
8	DCR	0.90	0.44	0.44	-0.28	-0.45		0.85%
9	DCR	0.89	0.57	0.57	-0.34	-0.49		0.97%
10	DCR	0.77	0.43	0.43	-0.25	-0.37		1.09%
11	CR	0.88	0.74	-0.51	-0.34	0.55		0.09%
12	CR	0.91	0.74	-0.51	-0.34	0.55		0.09%
13	CR	0.93	0.70	-0.48	-0.40	0.58		1.00%
14	CR	0.77	0.70	-0.54	-0.17	0.51		1.17%
15	CR	0.76	0.68	-0.52	-0.15	0.49		1.12%
16	CR	0.82	0.75	-0.52	-0.28	0.55		0.21%
17	CR	0.87	0.77	-0.50	-0.30	0.56		0.19%
18	CR	0.73	0.73	-0.55	-0.21	0.55		2.85%
19	CR	0.87	0.74	-0.48	-0.43	0.61		2.77%
20	CR	0.84	0.81	-0.50	-0.44	-0.19	0.58	0.40%

Table 26A. Item Analysis Grade Span 3: 6–8, Listening

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.37	0.26	-0.12	0.26	-0.22		0.11%
2	MC	0.92	0.44	0.44	-0.32	-0.30		0.08%
3	MC	0.85	0.42	-0.25	-0.32	0.42		0.08%
4	MC	0.82	0.44	-0.30	0.44	-0.31		0.09%
5	MC	0.45	0.28	-0.17	0.28	-0.17		0.11%
6	MC	0.94	0.43	0.43	-0.31	-0.29		0.08%
7	MC	0.80	0.44	0.44	-0.28	-0.30		0.11%
8	MC	0.54	0.34	-0.28	0.34	-0.14		0.13%
9	MC	0.88	0.44	-0.30	-0.30	0.44		0.11%
10	MC	0.79	0.31	0.31	-0.23	-0.20		0.13%
11	MC	0.93	0.39	-0.32	0.39	-0.21		0.05%
12	MC	0.90	0.41	-0.27	-0.30	0.41		0.09%
13	MC	0.75	0.39	-0.30	-0.24	0.39		0.07%
14	MC	0.87	0.39	-0.29	-0.29	0.39		0.09%
15	MC	0.84	0.43	0.43	-0.30	-0.30		0.11%
16	MC	0.85	0.46	-0.29	-0.33	0.46		0.09%
17	MC	0.75	0.37	0.37	-0.25	-0.26		0.11%
18	MC	0.79	0.43	-0.31	0.43	-0.26		0.13%
19	MC	0.59	0.39	-0.28	0.39	-0.24		0.14%
20	MC	0.82	0.54	-0.41	0.54	-0.31		0.13%
21	MC	0.76	0.39	-0.25	-0.32	0.39		0.11%
22	MC	0.91	0.40	0.40	-0.31	-0.26		0.14%
23	MC	0.92	0.42	-0.29	0.42	-0.30		0.15%

Table 27A. Item Analysis Grade Span 3: 6–8, Reading

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.71	0.27	0.27	-0.20	-0.15		0.10%
2	MC	0.93	0.48	-0.34	0.48	-0.32		0.04%
3	MC	0.80	0.38	0.38	-0.32	-0.18		0.08%
4	MC	0.81	0.44	0.44	-0.26	-0.33		0.06%
5	MC	0.63	0.28	-0.18	-0.17	0.28		0.09%
6	MC	0.61	0.27	-0.11	0.27	-0.24		0.10%
7	MC	0.87	0.45	-0.28	-0.33	0.45		0.13%
8	MC	0.74	0.49	-0.39	-0.29	0.49		0.11%
9	MC	0.63	0.37	-0.28	0.37	-0.18		0.16%
10	MC	0.80	0.39	0.39	-0.28	-0.25		0.20%
11	MC	0.51	0.31	0.31	-0.21	-0.16		0.11%
12	MC	0.89	0.43	-0.28	-0.30	0.43		0.12%
13	MC	0.70	0.41	0.41	-0.25	-0.27		0.15%
14	MC	0.75	0.42	-0.32	0.42	-0.24		0.15%
15	MC	0.69	0.45	0.45	-0.33	-0.27		0.13%
16	MC	0.73	0.42	0.42	-0.27	-0.32		0.16%
17	MC	0.88	0.40	-0.31	-0.26	0.40		0.15%
18	MC	0.56	0.39	-0.35	0.39	-0.10		0.20%
19	MC	0.41	0.29	-0.19	-0.12	0.29		0.20%
20	MC	0.84	0.50	-0.34	0.50	-0.34		0.18%
21	MC	0.72	0.50	-0.34	-0.25	0.50	-0.17	0.20%
22	MC	0.57	0.40	-0.22	-0.22	-0.15	0.41	0.19%
23	MC	0.51	0.32	-0.22	0.32	-0.09	-0.17	0.23%
24	MC	0.40	0.42	0.42	-0.06	-0.21	-0.28	0.30%
25	MC	0.42	0.21	0.06	-0.29	-0.22	0.21	0.20%
26	MC	0.87	0.51	-0.27	-0.33	0.51	-0.24	0.32%
27	MC	0.51	0.32	-0.25	-0.19	-0.06	0.32	0.40%
28	MC	0.75	0.52	-0.29	0.52	-0.28	-0.28	0.36%
29	MC	0.82	0.55	-0.34	0.55	-0.29	-0.25	0.34%
30	MC	0.70	0.46	0.46	-0.25	-0.21	-0.30	0.39%
31	MC	0.61	0.44	-0.29	-0.15	0.44	-0.21	0.42%
32	MC	0.59	0.46	-0.16	0.46	-0.28	-0.29	0.49%
33	MC	0.58	0.46	0.46	-0.26	-0.18	-0.23	0.54%
34	MC	0.62	0.40	0.40	-0.21	-0.23	-0.15	0.53%
35	MC	0.66	0.42	-0.14	-0.25	0.42	-0.24	0.48%

Table 28A. Item Analysis Grade Span 3: 6–8, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.83	0.37	0.37	-0.26	-0.26		0.05%
2	MC	0.91	0.45	-0.31	0.45	-0.31		0.05%
3	MC	0.85	0.51	0.51	-0.42	-0.28		0.08%
4	MC	0.94	0.37	-0.31	-0.20	0.37		0.04%
5	MC	0.80	0.38	-0.32	0.38	-0.20		0.11%
6	MC	0.80	0.36	-0.26	-0.24	0.36		0.23%
7	MC	0.69	0.38	0.38	-0.21	-0.42		0.15%
8	MC	0.72	0.39	-0.30	-0.27	0.39		0.14%
9	MC	0.77	0.45	-0.22	0.45	-0.39		0.13%
10	MC	0.51	0.35	0.35	-0.29	-0.11		0.13%
11	MC	0.87	0.44	-0.32	0.44	-0.32		0.10%
12	MC	0.92	0.50	-0.37	0.50	-0.33		0.08%
13	MC	0.88	0.51	0.51	-0.31	-0.40		0.12%
14	MC	0.88	0.45	-0.33	-0.30	0.45		0.11%
15	MC	0.30	0.07	0.07	-0.36	0.09		0.13%
16	MC	0.87	0.56	-0.36	-0.41	0.56		0.13%
17	MC	0.58	0.26	-0.08	0.26	-0.28		0.13%
18	MC	0.74	0.43	-0.32	-0.26	0.43		0.18%
19	MC	0.64	0.33	-0.32	-0.18	0.33		0.19%
20	MC	0.89	0.49	-0.34	0.49	-0.33		0.15%
21	CR	0.79	0.50	-0.30	-0.23	0.39		0.30%
22	CR	0.74	0.51	-0.35	-0.18	0.34		0.49%
23	CR	0.79	0.55	-0.36	-0.22	0.35		0.37%
24	CR	0.81	0.58	-0.38	-0.27	0.41		0.60%
25	CR	0.75	0.67	-0.19	-0.52	0.03	0.40	1.14%

Table 29A. Item Analysis Grade Span 3: 6–8, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.37	0.19	-0.07	0.19	-0.19		0.08%
2	MC	0.93	0.38	0.38	-0.27	-0.26		0.07%
3	MC	0.85	0.35	-0.21	-0.27	0.35		0.07%
4	MC	0.82	0.40	-0.26	0.40	-0.29		0.08%
5	MC	0.45	0.24	-0.15	0.24	-0.14		0.10%
6	MC	0.94	0.37	0.37	-0.26	-0.25		0.06%
7	MC	0.80	0.39	0.39	-0.24	-0.28		0.10%
8	MC	0.88	0.39	-0.28	-0.26	0.39		0.10%
9	MC	0.93	0.34	-0.29	0.34	-0.17		0.04%
10	MC	0.90	0.34	-0.23	-0.25	0.34		0.08%
11	MC	0.87	0.32	-0.23	-0.26	0.33		0.07%
12	MC	0.84	0.34	0.34	-0.23	-0.24		0.10%
13	MC	0.85	0.37	-0.24	-0.27	0.37		0.08%
14	MC	0.75	0.30	0.30	-0.20	-0.21		0.10%
15	MC	0.79	0.39	-0.28	0.39	-0.23		0.11%
16	MC	0.59	0.30	-0.24	0.30	-0.17		0.12%
17	MC	0.82	0.48	-0.38	0.48	-0.26		0.11%
18	MC	0.76	0.32	-0.20	-0.28	0.32		0.09%
19	MC	0.91	0.34	0.34	-0.27	-0.21		0.11%
20	MC	0.92	0.35	-0.24	0.35	-0.25		0.13%
21	MC	0.93	0.51	-0.36	0.51	-0.35		0.03%
22	MC	0.80	0.38	0.38	-0.32	-0.19		0.07%
23	MC	0.81	0.44	0.44	-0.26	-0.33		0.06%
24	MC	0.63	0.27	-0.18	-0.16	0.27		0.08%
25	MC	0.87	0.43	-0.28	-0.31	0.43		0.13%
26	MC	0.74	0.48	-0.41	-0.27	0.48		0.11%
27	MC	0.51	0.30	0.30	-0.20	-0.14		0.10%
28	MC	0.89	0.42	-0.28	-0.30	0.42		0.11%
29	MC	0.70	0.39	0.39	-0.24	-0.26		0.15%
30	MC	0.75	0.40	-0.32	0.40	-0.22		0.15%
31	MC	0.69	0.45	0.45	-0.32	-0.27		0.12%
32	MC	0.73	0.41	0.41	-0.27	-0.32		0.15%
33	MC	0.88	0.39	-0.30	-0.26	0.39		0.14%
34	MC	0.56	0.38	-0.35	0.38	-0.08		0.20%
35	MC	0.41	0.27	-0.18	-0.11	0.27		0.20%
36	MC	0.84	0.50	-0.34	0.50	-0.33		0.17%
37	MC	0.72	0.48	-0.33	-0.24	0.48	-0.16	0.18%
38	MC	0.57	0.39	-0.22	-0.20	-0.15	0.39	0.18%
39	MC	0.51	0.30	-0.22	0.30	-0.07	-0.16	0.22%
40	MC	0.40	0.40	0.40	-0.04	-0.22	-0.27	0.29%
41	MC	0.87	0.50	-0.27	-0.32	0.50	-0.25	0.30%
42	MC	0.51	0.30	-0.25	-0.19	-0.04	0.30	0.39%
43	MC	0.75	0.50	-0.28	0.50	-0.26	-0.29	0.35%
44	MC	0.82	0.53	-0.33	0.53	-0.29	-0.25	0.33%
45	MC	0.70	0.45	0.45	-0.24	-0.19	-0.30	0.38%

Table 29A. Item Analysis Grade Span 3: 6–8, Comprehension (continued)

Item Number	Item Type	*P*-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
46	MC	0.61	0.41	-0.27	-0.13	0.41	-0.19	0.41%
47	MC	0.60	0.43	-0.15	0.43	-0.26	-0.29	0.48%
48	MC	0.58	0.43	0.43	-0.24	-0.17	-0.23	0.53%
49	MC	0.62	0.38	0.38	-0.20	-0.22	-0.14	0.52%
50	MC	0.66	0.40	-0.13	-0.23	0.40	-0.24	0.47%

Table 30A. Item Analysis Grade Span 3: 6–8, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.82	0.55	0.55	-0.34	-0.42		0.11%
2	DCR	0.91	0.46	0.46	-0.31	-0.43		0.11%
3	DCR	0.31	0.33	0.33	-0.08	-0.26		0.30%
4	DCR	0.89	0.38	0.38	-0.27	-0.33		0.66%
5	DCR	0.74	0.46	0.46	-0.39	-0.24		0.82%
6	DCR	0.85	0.49	0.49	-0.41	-0.35		0.79%
7	DCR	0.76	0.49	0.49	-0.30	-0.37		0.86%
8	DCR	0.90	0.42	0.42	-0.26	-0.43		0.79%
9	DCR	0.89	0.54	0.55	-0.33	-0.47		0.92%
10	DCR	0.77	0.40	0.40	-0.24	-0.35		1.04%
11	CR	0.88	0.70	-0.49	-0.31	0.50		0.08%
12	CR	0.91	0.70	-0.47	-0.31	0.51		0.08%
13	CR	0.93	0.65	-0.45	-0.36	0.54		0.91%
14	CR	0.77	0.67	-0.51	-0.15	0.47		1.08%
15	CR	0.76	0.65	-0.50	-0.13	0.46		1.03%
16	CR	0.82	0.73	-0.50	-0.26	0.52		0.21%
17	CR	0.87	0.74	-0.48	-0.27	0.52		0.17%
18	CR	0.73	0.69	-0.51	-0.18	0.50		2.70%
19	CR	0.87	0.70	-0.45	-0.38	0.56		2.63%
20	CR	0.84	0.77	-0.49	-0.42	-0.16	0.54	0.38%
21	MC	0.37	0.12	-0.01	0.12	-0.18		0.11%
22	MC	0.92	0.43	0.43	-0.30	-0.29		0.08%
23	MC	0.85	0.31	-0.18	-0.25	0.31		0.08%
24	MC	0.82	0.36	-0.21	0.36	-0.29		0.09%
25	MC	0.46	0.17	-0.12	0.17	-0.09		0.11%
26	MC	0.94	0.43	0.43	-0.30	-0.31		0.08%
27	MC	0.80	0.37	0.37	-0.24	-0.25		0.11%
28	MC	0.54	0.22	-0.17	0.22	-0.10		0.12%
29	MC	0.88	0.40	-0.25	-0.30	0.40		0.11%
30	MC	0.79	0.22	0.22	-0.20	-0.12		0.12%
31	MC	0.93	0.36	-0.30	0.36	-0.20		0.04%
32	MC	0.90	0.38	-0.23	-0.30	0.38		0.08%
33	MC	0.75	0.28	-0.21	-0.21	0.28		0.06%
34	MC	0.87	0.30	-0.19	-0.29	0.30		0.08%
35	MC	0.84	0.33	0.33	-0.22	-0.24		0.11%
36	MC	0.85	0.35	-0.23	-0.25	0.35		0.09%
37	MC	0.75	0.28	0.28	-0.20	-0.19		0.10%
38	MC	0.79	0.34	-0.24	0.34	-0.22		0.12%
39	MC	0.59	0.31	-0.26	0.32	-0.18		0.13%
40	MC	0.82	0.53	-0.39	0.53	-0.31		0.13%
41	MC	0.76	0.30	-0.17	-0.30	0.30		0.10%
42	MC	0.91	0.37	0.37	-0.30	-0.22		0.13%
43	MC	0.92	0.40	-0.29	0.40	-0.28		0.15%

Table 31A. Item Analysis Grade Span 4: 9–12, Speaking

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.84	0.61	0.61	-0.35	-0.48		0.18%
2	DCR	0.90	0.63	0.63	-0.44	-0.47		0.15%
3	DCR	0.50	0.48	0.48	-0.19	-0.37		0.23%
4	DCR	0.87	0.57	0.57	-0.39	-0.45		1.23%
5	DCR	0.51	0.44	0.44	-0.16	-0.34		1.48%
6	DCR	0.75	0.55	0.55	-0.29	-0.45		1.58%
7	DCR	0.85	0.63	0.63	-0.37	-0.49		1.53%
8	DCR	0.89	0.56	0.56	-0.37	-0.45		1.52%
9	DCR	0.69	0.40	0.40	-0.19	-0.36		1.54%
10	DCR	0.91	0.52	0.52	-0.34	-0.43		1.63%
11	CR	0.88	0.79	-0.58	-0.35	0.61		0.27%
12	CR	0.90	0.80	-0.59	-0.39	0.67		0.21%
13	CR	0.78	0.68	-0.56	-0.19	0.49		1.15%
14	CR	0.77	0.79	-0.65	-0.17	0.57		1.27%
15	CR	0.82	0.78	-0.61	-0.27	0.59		1.44%
16	CR	0.89	0.84	-0.48	-0.44	0.63		0.44%
17	CR	0.85	0.81	-0.52	-0.38	0.60		0.39%
18	CR	0.86	0.77	-0.54	-0.37	0.62		2.81%
19	CR	0.87	0.75	-0.53	-0.39	0.61		2.79%
20	CR	0.84	0.85	-0.51	-0.51	-0.18	0.61	0.57%

Table 32A. Item Analysis Grade Span 4: 9–12, Listening

Item Number	Item Type	*P*-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.90	0.49	-0.31	0.49	-0.36		0.10%
2	MC	0.90	0.45	0.45	-0.35	-0.27		0.16%
3	MC	0.63	0.23	-0.13	-0.16	0.23		0.24%
4	MC	0.92	0.39	-0.23	-0.31	0.39		0.09%
5	MC	0.88	0.47	-0.38	-0.26	0.47		0.12%
6	MC	0.85	0.52	0.52	-0.41	-0.31		0.17%
7	MC	0.87	0.50	-0.29	0.50	-0.39		0.10%
8	MC	0.75	0.41	-0.20	0.41	-0.33		0.10%
9	MC	0.69	0.38	-0.26	-0.23	0.39		0.17%
10	MC	0.62	0.40	-0.27	0.40	-0.22		0.14%
11	MC	0.82	0.38	-0.28	0.38	-0.24		0.11%
12	MC	0.82	0.47	0.47	-0.33	-0.31		0.11%
13	MC	0.73	0.50	-0.32	-0.34	0.50		0.14%
14	MC	0.60	0.38	-0.29	0.39	-0.21		0.16%
15	MC	0.66	0.42	0.42	-0.26	-0.29		0.15%
16	MC	0.67	0.41	-0.25	-0.26	0.41		0.14%
17	MC	0.76	0.35	-0.28	-0.21	0.35		0.16%
18	MC	0.81	0.43	0.43	-0.30	-0.27		0.15%
19	MC	0.80	0.43	-0.27	-0.30	0.43		0.22%
20	MC	0.90	0.40	0.40	-0.34	-0.21		0.15%
21	MC	0.54	0.38	0.38	-0.32	-0.17		0.22%
22	MC	0.83	0.51	-0.35	-0.35	0.52		0.18%
23	MC	0.64	0.42	-0.15	0.42	-0.36		0.26%

Table 33A. Item Analysis Grade Span 4: 9–12, Reading

Item Number	Item Type	*P*-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.53	0.31	-0.22	-0.16	0.31		0.15%
2	MC	0.58	0.23	-0.08	0.23	-0.20		0.06%
3	MC	0.90	0.48	-0.31	0.48	-0.36		0.06%
4	MC	0.51	0.19	0.19	-0.20	-0.04		0.07%
5	MC	0.87	0.54	-0.38	-0.36	0.54		0.08%
6	MC	0.64	0.35	-0.20	0.35	-0.26		0.14%
7	MC	0.78	0.50	-0.30	-0.37	0.50		0.11%
8	MC	0.79	0.43	0.43	-0.30	-0.27		0.13%
9	MC	0.65	0.47	-0.38	0.47	-0.22		0.16%
10	MC	0.61	0.32	-0.11	0.32	-0.27		0.22%
11	MC	0.50	0.34	0.34	-0.13	-0.32		0.22%
12	MC	0.79	0.46	0.46	-0.34	-0.27		0.16%
13	MC	0.83	0.43	-0.27	0.43	-0.31		0.14%
14	MC	0.79	0.50	-0.29	-0.37	0.50		0.20%
15	MC	0.86	0.45	-0.34	0.45	-0.26		0.09%
16	MC	0.70	0.49	-0.38	0.49	-0.24		0.18%
17	MC	0.69	0.42	0.42	-0.23	-0.31		0.18%
18	MC	0.46	0.30	-0.09	-0.29	0.30		0.21%
19	MC	0.52	0.44	-0.29	-0.22	0.44		0.16%
20	MC	0.65	0.50	0.50	-0.33	-0.32		0.18%
21	MC	0.49	0.20	0.20	-0.20	-0.24	0.10	0.21%
22	MC	0.36	0.31	-0.10	-0.16	-0.11	0.31	0.14%
23	MC	0.49	0.28	0.28	-0.07	-0.18	-0.15	0.17%
24	MC	0.58	0.49	-0.24	-0.29	-0.19	0.49	0.22%
25	MC	0.60	0.39	-0.25	0.39	-0.16	-0.15	0.25%
26	MC	0.67	0.57	-0.26	-0.33	-0.28	0.57	0.42%
27	MC	0.62	0.48	-0.28	0.48	-0.20	-0.25	0.51%
28	MC	0.67	0.51	0.51	-0.24	-0.29	-0.25	0.58%
29	MC	0.74	0.54	-0.33	-0.24	0.54	-0.28	0.71%
30	MC	0.38	0.26	-0.05	0.26	-0.09	-0.22	0.77%
31	MC	0.36	0.37	0.37	-0.14	-0.26	-0.07	0.89%
32	MC	0.67	0.48	-0.26	0.48	-0.20	-0.27	1.08%
33	MC	0.70	0.47	-0.26	-0.25	0.47	-0.22	1.19%
34	MC	0.65	0.47	-0.20	-0.20	-0.30	0.47	1.28%
35	MC	0.56	0.46	-0.15	-0.28	0.46	-0.25	1.28%

Table 34A. Item Analysis Grade Span 4: 9–12, Writing

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.58	0.38	-0.34	-0.24	0.39		0.20%
2	MC	0.88	0.54	-0.28	-0.45	0.54		0.02%
3	MC	0.90	0.54	-0.35	-0.40	0.54		0.06%
4	MC	0.92	0.53	-0.43	0.53	-0.29		0.06%
5	MC	0.89	0.56	0.56	-0.37	-0.41		0.07%
6	MC	0.83	0.56	-0.49	0.56	-0.25		0.06%
7	MC	0.81	0.52	0.52	-0.29	-0.43		0.19%
8	MC	0.87	0.61	-0.52	0.61	-0.29		0.10%
9	MC	0.87	0.49	0.49	-0.31	-0.39		0.09%
10	MC	0.50	0.17	0.17	-0.15	-0.07		0.09%
11	MC	0.91	0.49	-0.29	0.49	-0.38		0.08%
12	MC	0.68	0.43	0.43	-0.26	-0.35		0.14%
13	MC	0.88	0.49	-0.35	-0.32	0.49		0.16%
14	MC	0.70	0.47	0.47	-0.27	-0.34		0.14%
15	MC	0.58	0.33	-0.34	-0.14	0.33		0.27%
16	MC	0.74	0.46	0.46	-0.33	-0.28		0.22%
17	MC	0.85	0.38	-0.25	0.38	-0.29		0.27%
18	MC	0.52	0.29	-0.24	-0.11	0.29		0.36%
19	MC	0.49	0.28	0.28	-0.11	-0.21		0.48%
20	MC	0.66	0.43	-0.36	-0.21	0.43		0.37%
21	CR	0.81	0.57	-0.36	-0.37	0.47		0.75%
22	CR	0.82	0.60	-0.38	-0.38	0.51		1.03%
23	CR	0.78	0.59	-0.35	-0.32	0.46		1.42%
24	CR	0.79	0.57	-0.35	-0.31	0.43		1.34%
25	CR	0.73	0.70	-0.28	-0.55	0.16	0.37	1.72%

Table 35A. Item Analysis Grade Span 4: 9–12, Comprehension

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	MC	0.90	0.43	-0.27	0.43	-0.33		0.09%
2	MC	0.90	0.39	0.39	-0.31	-0.24		0.14%
3	MC	0.92	0.35	-0.20	-0.28	0.35		0.06%
4	MC	0.88	0.42	-0.33	-0.23	0.42		0.10%
5	MC	0.86	0.46	0.46	-0.36	-0.28		0.14%
6	MC	0.87	0.45	-0.26	0.45	-0.36		0.07%
7	MC	0.75	0.37	-0.18	0.37	-0.30		0.08%
8	MC	0.62	0.37	-0.26	0.37	-0.20		0.11%
9	MC	0.82	0.33	-0.23	0.33	-0.23		0.10%
10	MC	0.82	0.41	0.41	-0.30	-0.26		0.08%
11	MC	0.73	0.44	-0.28	-0.31	0.44		0.10%
12	MC	0.60	0.33	-0.27	0.33	-0.17		0.12%
13	MC	0.67	0.36	0.36	-0.24	-0.24		0.11%
14	MC	0.76	0.31	-0.26	-0.18	0.31		0.12%
15	MC	0.81	0.40	0.40	-0.29	-0.25		0.11%
16	MC	0.80	0.39	-0.24	-0.28	0.39		0.19%
17	MC	0.90	0.35	0.35	-0.30	-0.18		0.10%
18	MC	0.54	0.34	0.34	-0.30	-0.14		0.17%
19	MC	0.83	0.47	-0.31	-0.32	0.47		0.13%
20	MC	0.64	0.36	-0.11	0.36	-0.33		0.20%
21	MC	0.90	0.51	-0.31	0.51	-0.39		0.03%
22	MC	0.87	0.57	-0.40	-0.38	0.57		0.05%
23	MC	0.64	0.32	-0.17	0.32	-0.26		0.10%
24	MC	0.78	0.49	-0.28	-0.36	0.49		0.08%
25	MC	0.79	0.41	0.41	-0.28	-0.26		0.10%
26	MC	0.65	0.48	-0.38	0.48	-0.23		0.13%
27	MC	0.50	0.33	0.33	-0.11	-0.34		0.19%
28	MC	0.79	0.45	0.45	-0.33	-0.27		0.14%
29	MC	0.83	0.43	-0.26	0.43	-0.32		0.11%
30	MC	0.79	0.52	-0.30	-0.38	0.52		0.18%
31	MC	0.86	0.47	-0.36	0.47	-0.28		0.06%
32	MC	0.70	0.50	-0.38	0.50	-0.25		0.16%
33	MC	0.46	0.27	-0.07	-0.28	0.28		0.18%
34	MC	0.52	0.41	-0.27	-0.21	0.42		0.14%
35	MC	0.65	0.49	0.49	-0.32	-0.31		0.15%
36	MC	0.49	0.18	0.18	-0.19	-0.24	0.11	0.19%
37	MC	0.36	0.29	-0.08	-0.16	-0.10	0.29	0.13%
38	MC	0.49	0.25	0.25	-0.06	-0.17	-0.13	0.15%
39	MC	0.58	0.47	-0.24	-0.27	-0.17	0.47	0.21%
40	MC	0.60	0.36	-0.24	0.36	-0.14	-0.14	0.23%
41	MC	0.67	0.55	-0.26	-0.32	-0.27	0.55	0.41%
42	MC	0.62	0.46	-0.28	0.46	-0.19	-0.25	0.48%
43	MC	0.67	0.49	0.49	-0.24	-0.28	-0.25	0.56%
44	MC	0.74	0.53	-0.34	-0.22	0.53	-0.28	0.68%
45	MC	0.38	0.24	-0.04	0.24	-0.08	-0.22	0.74%

Table 35A. Item Analysis Grade Span 4: 9–12, Comprehension (continued)

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
46	MC	0.36	0.34	0.34	-0.13	-0.25	-0.06	0.87%
47	MC	0.67	0.48	-0.27	0.48	-0.19	-0.28	1.06%
48	MC	0.70	0.45	-0.25	-0.24	0.45	-0.20	1.17%
49	MC	0.66	0.45	-0.20	-0.20	-0.29	0.45	1.26%
50	MC	0.56	0.44	-0.14	-0.27	0.44	-0.25	1.26%

Table 36A. Item Analysis Grade Span 4: 9–12, Oral

Item Number	Item Type	P-Value	Item-Total Correlation	Point Biserial				Percent Omit
				at A	at B	at C	at D	
1	DCR	0.85	0.60	0.60	-0.36	-0.46		0.19%
2	DCR	0.90	0.60	0.60	-0.43	-0.44		0.15%
3	DCR	0.51	0.49	0.49	-0.20	-0.37		0.23%
4	DCR	0.88	0.53	0.54	-0.37	-0.42		1.21%
5	DCR	0.51	0.45	0.45	-0.17	-0.35		1.44%
6	DCR	0.76	0.53	0.53	-0.28	-0.43		1.52%
7	DCR	0.85	0.62	0.62	-0.37	-0.48		1.48%
8	DCR	0.89	0.52	0.52	-0.35	-0.42		1.48%
9	DCR	0.69	0.41	0.41	-0.20	-0.36		1.49%
10	DCR	0.91	0.49	0.49	-0.32	-0.40		1.59%
11	CR	0.88	0.74	-0.55	-0.33	0.57		0.27%
12	CR	0.91	0.76	-0.56	-0.39	0.64		0.20%
13	CR	0.78	0.64	-0.52	-0.18	0.46		1.09%
14	CR	0.77	0.77	-0.63	-0.16	0.54		1.19%
15	CR	0.82	0.75	-0.58	-0.26	0.56		1.39%
16	CR	0.89	0.81	-0.46	-0.42	0.60		0.41%
17	CR	0.85	0.79	-0.49	-0.36	0.57		0.33%
18	CR	0.86	0.73	-0.51	-0.34	0.57		2.71%
19	CR	0.87	0.71	-0.50	-0.37	0.58		2.69%
20	CR	0.84	0.82	-0.48	-0.49	-0.17	0.59	0.50%
21	MC	0.90	0.47	-0.27	0.47	-0.37		0.09%
22	MC	0.90	0.41	0.42	-0.35	-0.23		0.15%
23	MC	0.63	0.14	-0.08	-0.10	0.14		0.23%
24	MC	0.92	0.37	-0.20	-0.31	0.37		0.07%
25	MC	0.88	0.45	-0.36	-0.26	0.45		0.11%
26	MC	0.86	0.50	0.50	-0.38	-0.33		0.16%
27	MC	0.88	0.49	-0.28	0.49	-0.39		0.09%
28	MC	0.75	0.31	-0.14	0.31	-0.27		0.10%
29	MC	0.69	0.26	-0.22	-0.13	0.26		0.16%
30	MC	0.62	0.30	-0.21	0.30	-0.16		0.14%
31	MC	0.83	0.31	-0.21	0.31	-0.21		0.11%
32	MC	0.82	0.41	0.41	-0.28	-0.27		0.11%
33	MC	0.73	0.43	-0.27	-0.29	0.43		0.12%
34	MC	0.60	0.30	-0.24	0.30	-0.16		0.15%
35	MC	0.67	0.38	0.38	-0.23	-0.26		0.15%
36	MC	0.67	0.29	-0.18	-0.19	0.29		0.13%
37	MC	0.76	0.25	-0.26	-0.11	0.25		0.15%
38	MC	0.81	0.32	0.32	-0.23	-0.20		0.14%
39	MC	0.80	0.35	-0.22	-0.25	0.35		0.21%
40	MC	0.90	0.37	0.37	-0.33	-0.18		0.13%
41	MC	0.54	0.30	0.30	-0.32	-0.09		0.20%
42	MC	0.83	0.46	-0.32	-0.30	0.46		0.17%
43	MC	0.64	0.34	-0.08	0.34	-0.34		0.23%

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix B: Comparison of 2008 and 2009/2010 Anchor Parameters
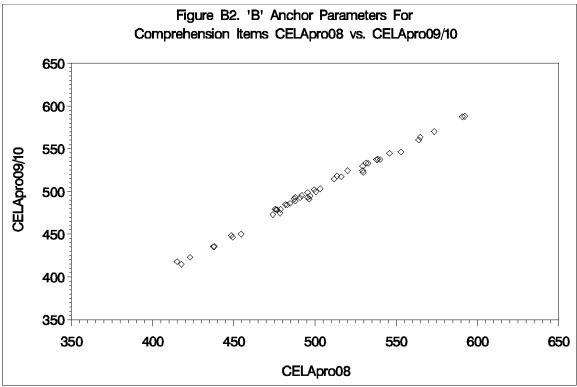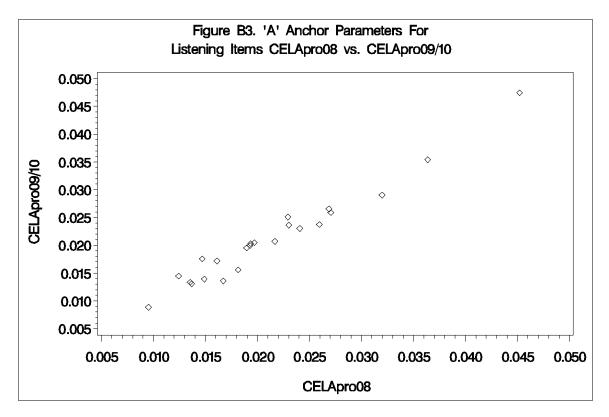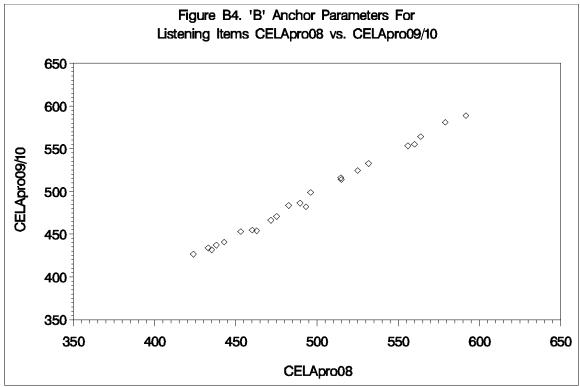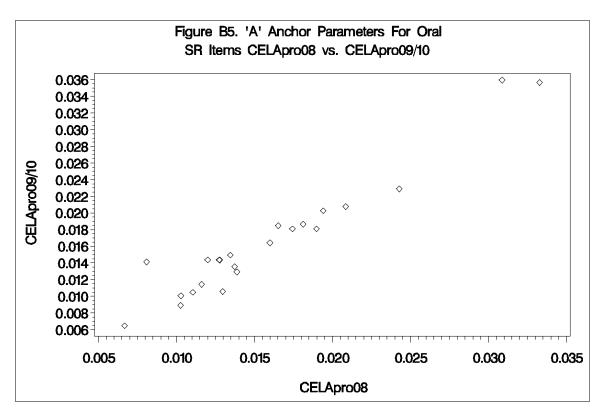
List of Tables

List of Figures

Table 1B. Comprehension Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10

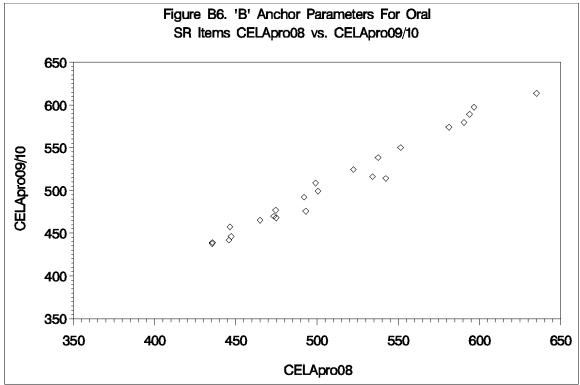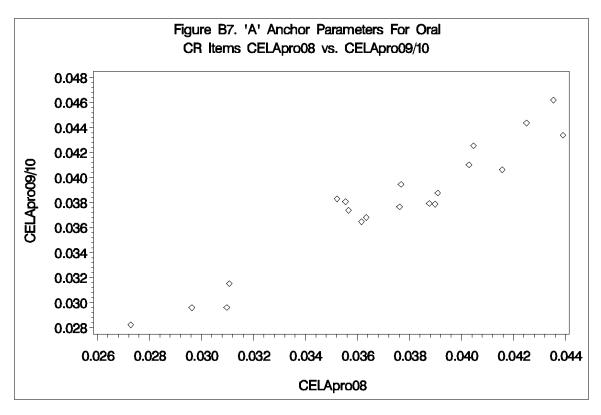| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	1	0.000	-0.006
CO	2	0.001	2.997
CO	3	-0.000	-2.862
CO	4	0.000	-3.358
CO	5	0.002	-3.519
CO	6	-0.001	4.058
CO	7	0.000	0.300
CO	8	-0.000	-2.833
CO	9	0.001	-1.562
CO	10	-0.000	-4.445
CO	11	-0.001	4.511
CO	12	-0.001	-1.910
CO	13	-0.000	-7.202
CO	14	-0.001	-2.504
CO	15	-0.000	-3.892
CO	16	-0.001	0.196
CO	17	0.008	-4.317
CO	18	-0.002	-4.698
CO	19	-0.001	-0.630
CO	21	-0.000	-3.277
CO	22	-0.000	2.577
CO	23	0.001	1.278
CO	10	0.001	-6.533
CO	11	-0.001	-5.407
CO	12	-0.000	3.631
CO	13	0.001	-0.690
CO	14	0.000	-2.559
CO	15	-0.000	1.535
CO	16	0.001	0.202
CO	17	0.001	2.034
CO	18	0.000	2.905
CO	19	-0.002	-2.083
CO	20	0.001	-2.287
CO	21	0.001	4.012

Table 1B. Comprehension Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	22	0.003	0.224
CO	23	0.000	1.09/106
CO	24	-0.001	-1.087
CO	25	0.001	2.818
CO	26	0.001	4.266
CO	27	0.002	3.494
CO	28	0.001	-1.124
CO	29	0.002	5.503
CO	30	0.003	-0.777
CO	31	0.001	1.930
CO	32	0.001	1.395
CO	33	-0.000	-0.104
CO	34	0.001	2.555
CO	35	0.002	2.223

Figure B1. 'A' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10

Figure B2. 'B' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10
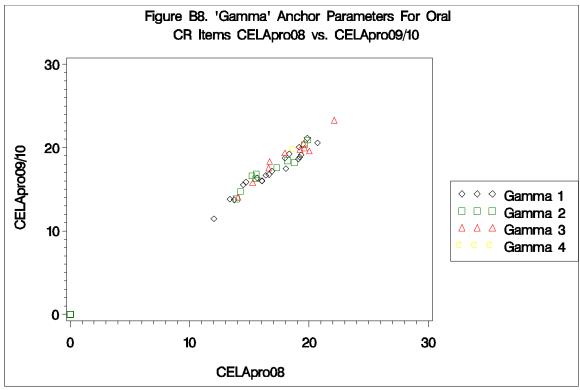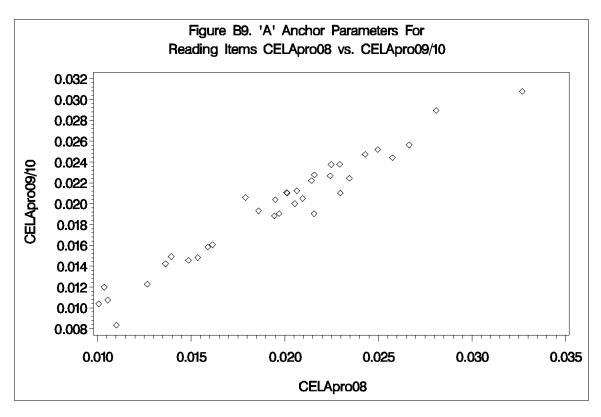
Table 2B. Listening Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
LI	1	-0.001	-3.472
LI	2	0.002	2.852
LI	3	0.001	0.971
LI	4	0.001	-2.858
LI	5	0.003	-2.468
LI	6	-0.003	-3.214
LI	7	0.001	0.147
LI	8	-0.003	-5.210
LI	9	0.001	0.412
LI	10	-0.002	-8.934
LI	11	-0.003	-1.014
LI	12	-0.001	-0.768
LI	13	-0.001	-11.19
LI	14	-0.001	-1.906
LI	15	-0.001	-4.453
LI	16	-0.001	1.058
LI	17	0.002	1.962
LI	18	-0.001	-0.516
LI	19	-0.000	1.138
LI	20	0.001	2.686
LI	21	0.001	-4.704
LI	22	-0.000	-5.376
LI	23	0.002	1.452

Figure B3. 'A' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Figure B4. 'B' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Table 3B. Oral Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
OR	1	0.005	10.902
OR	2	0.002	3.601
OR	3	0.000	2.398
OR	4	0.006	-21.53
OR	5	0.002	-7.257
OR	6	-0.000	-1.569
OR	7	0.001	0.109/10
OR	8	-0.001	-6.789
OR	9	0.002	-4.888
OR	10	0.001	-3.488
OR	11	-0.001	1.136
OR	12	-0.001	-3.756
OR	13	-0.001	-28.24
OR	14	-0.000	-0.788
OR	15	-0.001	-17.17
OR	16	-0.000	0.133
OR	17	-0.000	0.641
OR	18	-0.002	-17.88
OR	19	0.001	-1.313
OR	20	0.002	9.645
OR	21	0.002	-10.95
OR	22	-0.000	2.157
OR	23	0.002	2.09/107

Figure B5. 'A' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10

Figure B6. 'B' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10
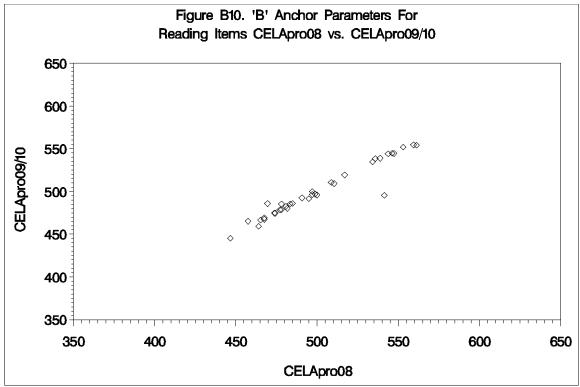
Table 4B. Oral Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (CR Items)

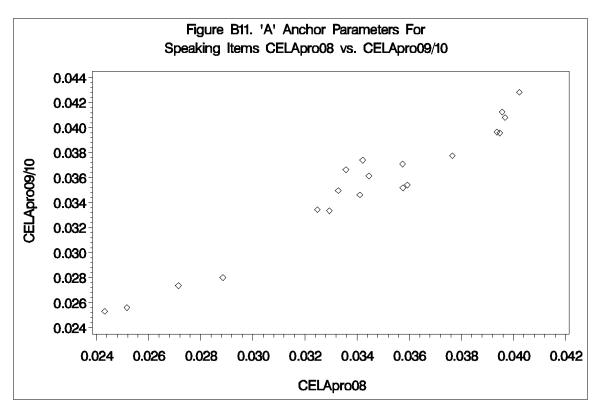| Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
OR	1	-0.001	-0.554	0.000		
OR	2	-0.001	-0.358	0.000		
OR	3	0.002	0.895	0.000		
OR	4	-0.000	-0.198	0.000		
OR	5	0.000	-0.064	0.000		
OR	6	0.003	1.281	0.000		
OR	7	-0.001	-0.123	0.000		
OR	8	0.000	-0.011	0.000		
OR	9	0.000	0.307	0.000		
OR	10	-0.001	-0.475	0.000		
OR	11	0.003	1.024	1.252	1.417	
OR	12	0.003	1.146	1.423	1.637	
OR	13	0.001	0.274	0.294	0.566	
OR	14	0.000	0.123	0.215	0.322	
OR	15	0.002	0.780	1.021	1.200	
OR	16	-0.000	-0.008	-0.058	0.017	
OR	17	0.002	0.896	0.741	0.986	
OR	18	-0.001	-0.560	-0.557	-0.431	
OR	19	0.001	0.460	0.463	0.547	
OR	20	0.002	0.703	0.732	0.951	1.029

Figure B7. 'A' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

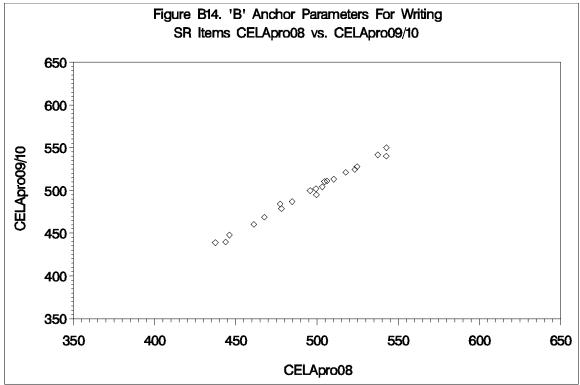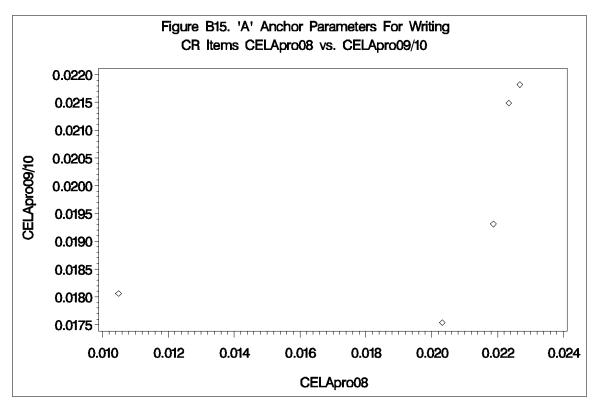Figure B8. 'Gamma' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

Table 5B. Reading Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	1	0.002	16.217
RD	2	0.001	1.714
RD	3	-0.001	1.596
RD	4	-0.000	-1.499
RD	5	0.001	-1.162
RD	6	-0.003	-45.76
RD	7	-0.001	0.031
RD	8	-0.000	-1.164
RD	9	0.001	-4.804
RD	10	0.001	-7.089
RD	11	-0.000	1.096
RD	12	-0.001	1.703
RD	13	0.000	-1.895
RD	14	-0.001	-2.780
RD	15	-0.001	0.706
RD	16	-0.000	-0.103
RD	17	-0.001	-1.695
RD	18	-0.001	0.874
RD	19	-0.003	-3.316
RD	20	0.000	-0.732
RD	21	0.001	7.490
RD	22	0.001	0.522
RD	23	0.000	1.218
RD	24	-0.002	-4.671
RD	25	0.001	3.091
RD	26	0.000	2.347
RD	27	0.001	2.137
RD	28	0.001	-1.124
RD	29	0.001	7.106
RD	30	0.003	0.474
RD	31	0.001	2.668
RD	32	0.000	0.405
RD	33	-0.002	-4.154

Table 5B. Reading Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	34	-0.000	1.957
RD	35	0.001	1.297

Figure B9. 'A' Anchor Parameters For Reading Items CELApro08 vs. CELApro09/10

Figure B10. 'B' Anchor Parameters For Reading Items CELApro08 vs. CELApro09/10

Table 6B. Speaking Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
SP	1	-0.001	-0.325
SP	2	0.000	0.130
SP	3	0.002	0.818
SP	4	0.000	-0.026
SP	5	0.000	0.106
SP	6	0.003	1.234
SP	7	0.000	0.247
SP	8	0.001	0.061
SP	9	0.000	0.285
SP	10	-0.001	-0.379
SP	11	0.003	1.285
SP	12	0.003	1.159
SP	13	0.001	0.437
SP	14	0.001	0.348
SP	15	0.001	0.423
SP	16	0.000	0.083
SP	17	0.002	0.791
SP	18	-0.001	-0.315
SP	19	0.001	0.483
SP	20	0.002	0.703

Figure B11. 'A' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10

Figure B12. 'B' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10
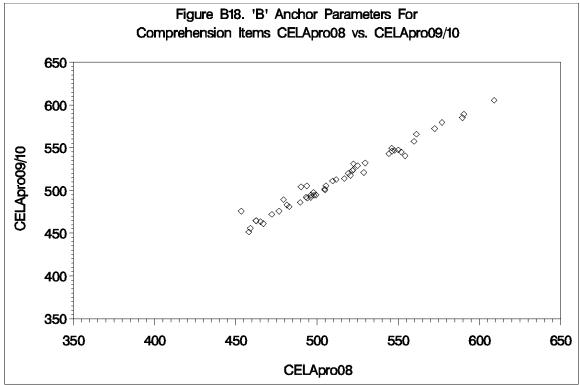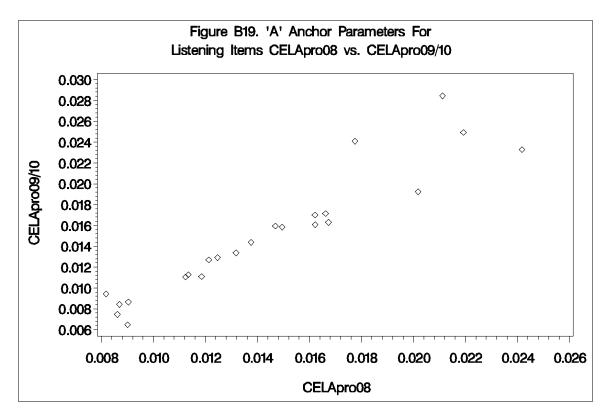
Table 7B. Writing Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
WR	1	-0.001	6.927
WR	2	-0.001	-3.889
WR	3	-0.002	4.098
WR	4	-0.001	1.019
WR	5	-0.001	1.531
WR	6	-0.002	1.681
WR	7	-0.000	2.893
WR	8	-0.001	1.195
WR	9	-0.001	-2.575
WR	10	-0.001	5.817
WR	11	-0.002	3.341
WR	12	0.000	7.147
WR	13	-0.002	-4.645
WR	14	-0.001	-0.650
WR	15	-0.002	4.099
WR	16	-0.001	3.585
WR	17	-0.002	1.232
WR	18	-0.001	4.798
WR	19	-0.004	0.736
WR	20	-0.004	2.631
WR	21	-0.005	2.077

Figure B13. 'A' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Figure B14. 'B' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Table 8B. Writing Grade Span 2: 3–5 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
WR	22	-0.003	-1.400	-1.086	-1.453	
WR	23	-0.003	-1.418	-0.872	-1.307	
WR	24	-0.001	-0.057	-0.396	-0.260	
WR	25	-0.001	-0.146	-0.376	-0.308	
WR	26	0.008	1.912	4.283	3.676	2.694

Figure B15. 'A' Anchor Parameters For Writing
CR Items CELApro08 vs. CELApro09/10

Figure B16. 'Gamma' Anchor Parameters For Writing
CR Items CELApro08 vs. CELApro09/10

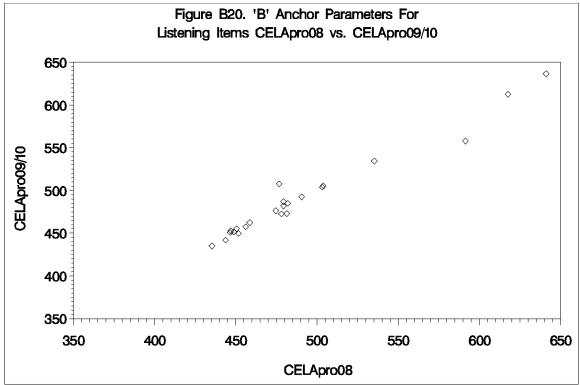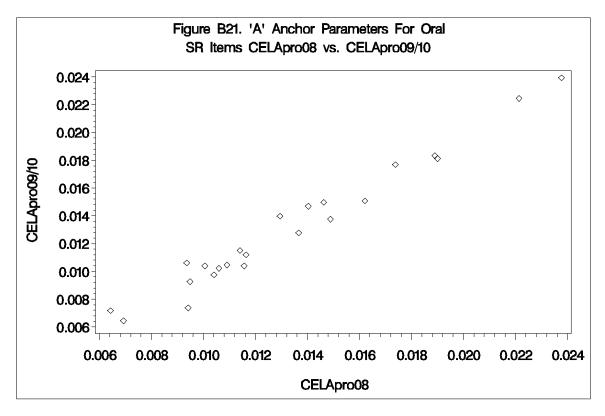Table 9B. Comprehension Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10

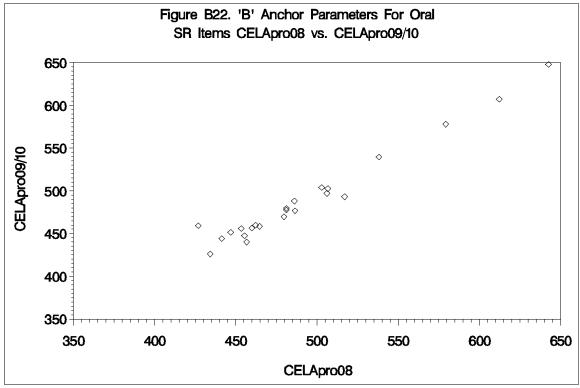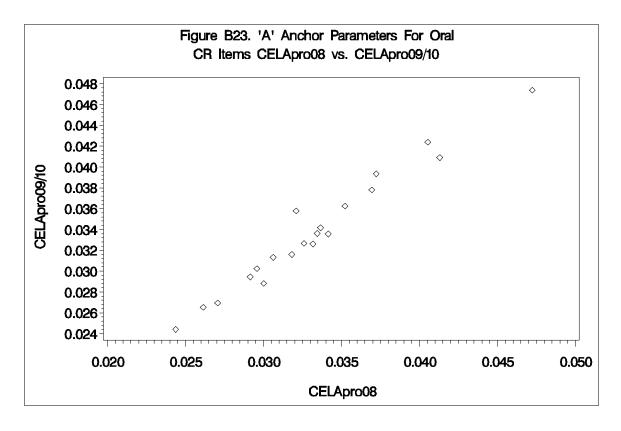| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	1	0.004	-3.597
CO	2	-0.002	-5.581
CO	3	0.002	22.434
CO	4	0.000	-0.210
CO	5	0.003	-4.424
CO	6	0.002	2.270
CO	7	-0.001	-4.723
CO	9	0.000	-0.833
CO	11	-0.002	-6.601
CO	12	0.000	-1.928
CO	14	0.000	0.095
CO	15	-0.001	-3.696
CO	16	0.000	-1.927
CO	17	0.000	-3.018
CO	18	0.000	11.698
CO	19	-0.002	-13.71
CO	20	-0.001	-1.625
CO	21	-0.001	-4.195
CO	22	0.002	2.030
CO	23	-0.000	-3.251
CO	2	0.004	1.654
CO	3	0.003	14.111
CO	4	-0.000	-2.836
CO	5	0.001	8.737
CO	7	-0.001	-4.195
CO	8	-0.001	-2.835
CO	11	0.003	2.558
CO	12	-0.001	-3.657
CO	13	-0.003	-7.808
CO	14	-0.000	1.589
CO	15	0.001	2.332
CO	16	-0.002	-2.762
CO	17	0.003	10.118
CO	18	-0.001	-6.880

Table 9B. Comprehension Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	19	-0.001	-1.221
CO	20	0.001	-1.090
CO	21	0.001	1.835
CO	22	-0.001	-2.559
CO	23	0.000	-2.067
CO	24	0.001	0.045
CO	26	0.001	-0.133
CO	27	0.001	4.607
CO	28	-0.000	0.571
CO	29	0.001	1.033
CO	30	0.003	1.783
CO	31	0.002	0.045
CO	32	0.004	3.762
CO	33	0.002	-0.772
CO	34	-0.001	-1.276
CO	35	0.001	4.173

Figure B17. 'A' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10

Figure B18. 'B' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10
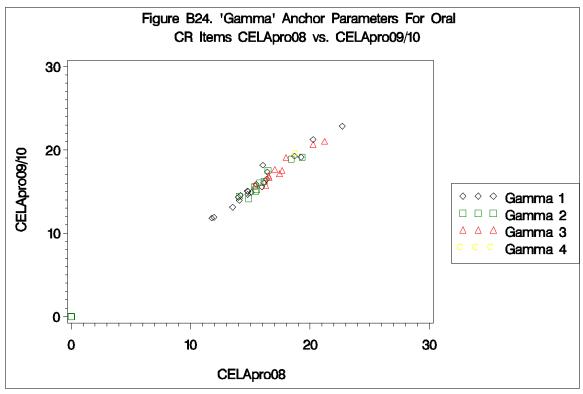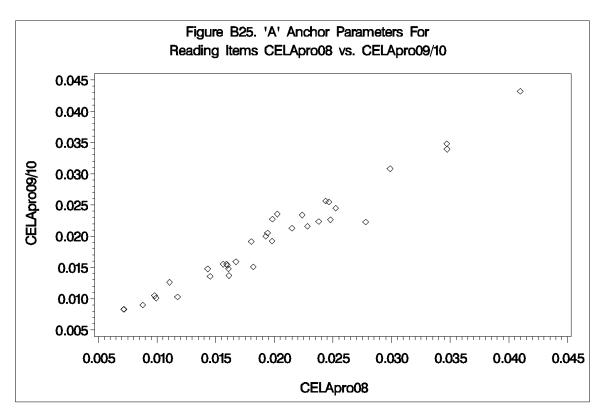
Table 10B. Listening Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
LI	1	0.006	-4.456
LI	2	-0.001	-2.015
LI	3	0.000	1.564
LI	4	0.000	2.628
LI	5	0.007	-5.015
LI	6	0.003	5.932
LI	7	-0.001	-5.215
LI	8	-0.003	-33.47
LI	9	0.001	3.657
LI	10	-0.001	-8.620
LI	11	-0.000	-0.428
LI	12	0.001	2.562
LI	13	-0.000	2.025
LI	14	0.001	5.017
LI	15	0.001	1.738
LI	16	0.001	3.156
LI	17	0.001	30.978
LI	18	-0.000	7.475
LI	19	-0.000	-0.916
LI	20	-0.001	0.696
LI	21	-0.000	1.818
LI	22	0.001	4.495
LI	23	-0.000	-1.861

Figure B19. 'A' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Figure B20. 'B' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Table 11B. Oral Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
OR	1	-0.001	5.347
OR	2	0.000	4.616
OR	3	0.001	32.335
OR	4	-0.000	-3.529
OR	5	-0.001	-4.914
OR	6	0.000	-3.274
OR	7	-0.001	-9.804
OR	8	0.001	-1.358
OR	9	0.001	-2.273
OR	10	-0.000	-6.295
OR	11	-0.001	-8.382
OR	12	0.000	-7.856
OR	13	-0.002	-23.95
OR	14	0.000	2.707
OR	15	0.000	-1.607
OR	16	-0.001	-9.977
OR	17	-0.000	-4.090
OR	18	-0.000	1.971
OR	19	-0.000	1.540
OR	20	0.000	1.137
OR	21	-0.001	-9.096
OR	22	0.001	2.849
OR	23	-0.001	-16.570

Figure B21. 'A' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10

Figure B22. 'B' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10
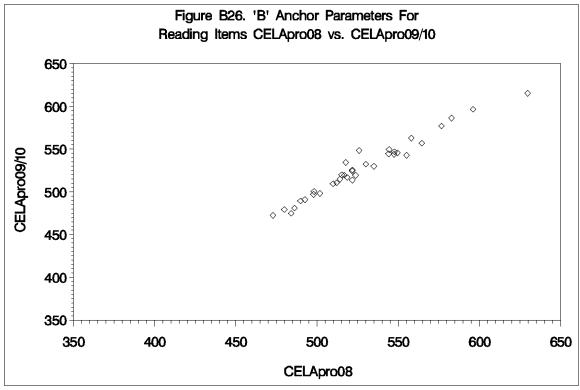
Table 12B. Oral Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (CR Items)

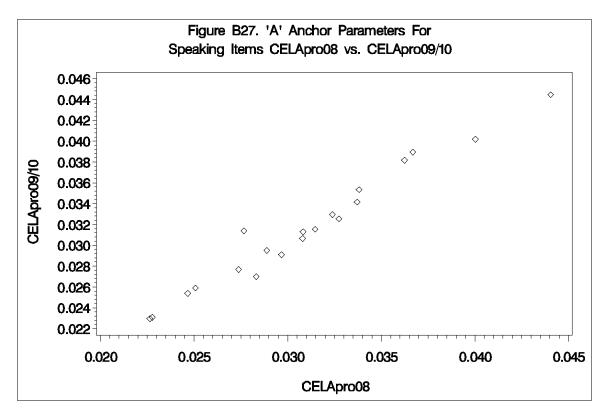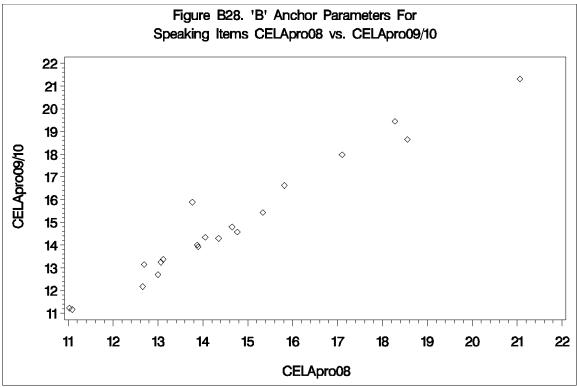| Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
OR	1	0.002	0.991	0.000		
OR	2	0.000	-0.135	0.000		
OR	3	-0.000	-0.090	0.000		
OR	4	0.000	0.043	0.000		
OR	5	0.000	0.196	0.000		
OR	6	-0.001	-0.394	0.000		
OR	7	0.004	2.132	0.000		
OR	8	0.001	0.312	0.000		
OR	9	0.000	0.150	0.000		
OR	10	0.000	-0.055	0.000		
OR	11	-0.001	-0.178	-0.378	-0.305	
OR	12	0.001	0.301	0.266	0.349	
OR	13	0.002	0.917	1.062	1.112	
OR	14	-0.001	-0.395	-0.717	-0.577	
OR	15	-0.000	-0.138	-0.302	-0.108	
OR	16	0.001	0.363	0.391	0.351	
OR	17	0.000	0.140	0.062	0.130	
OR	18	0.001	0.565	0.455	0.412	
OR	19	-0.000	-0.142	-0.226	-0.169	
OR	20	0.001	0.398	0.264	0.629	0.632

Figure B23. 'A' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

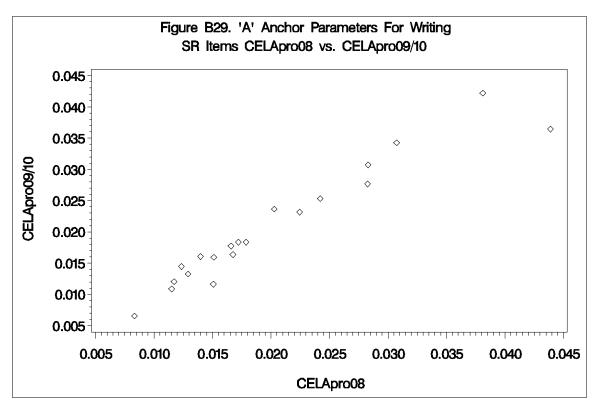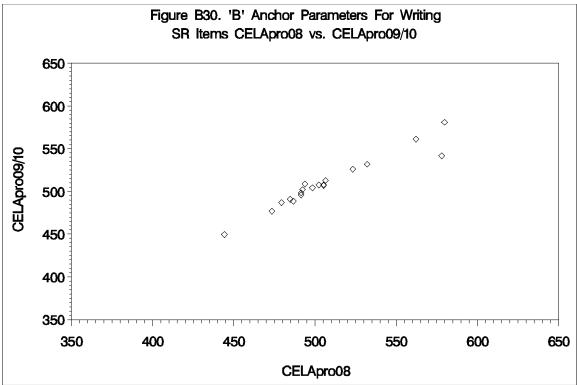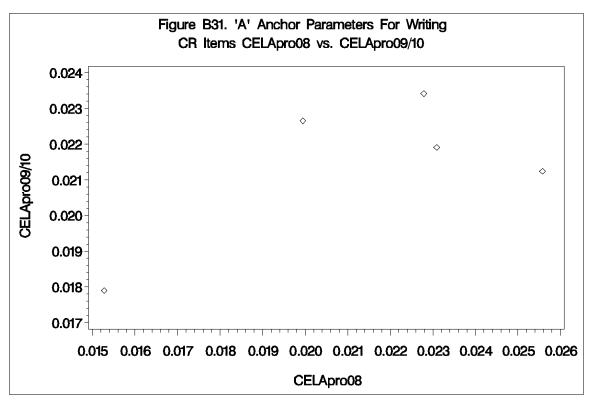Figure B24. 'Gamma' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

Table 13B. Reading Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10

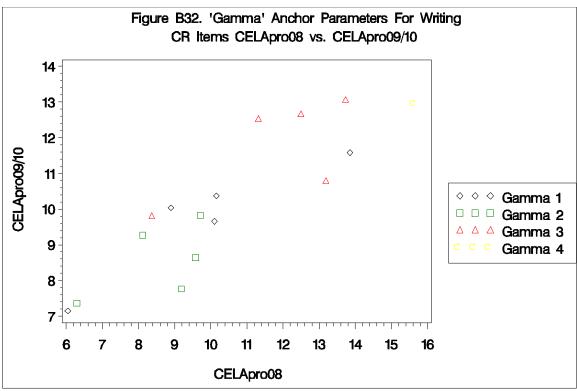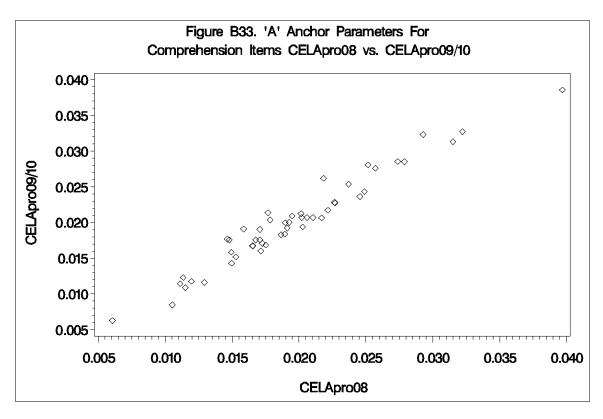| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	1	0.000	4.125
RD	2	0.002	-0.427
RD	3	-0.000	-1.667
RD	4	-0.001	-3.718
RD	5	0.000	-4.068
RD	6	0.001	22.487
RD	7	-0.001	-5.289
RD	8	-0.001	-4.443
RD	9	0.002	16.830
RD	10	-0.000	2.221
RD	11	0.001	3.609
RD	12	-0.002	-9.114
RD	13	-0.002	-7.990
RD	14	-0.001	-1.480
RD	15	0.001	2.232
RD	16	-0.001	2.568
RD	17	0.001	-0.464
RD	18	-0.003	-12.47
RD	19	-0.006	0.484
RD	20	0.001	-0.537
RD	21	-0.000	-1.546
RD	22	-0.001	-3.567
RD	23	0.001	4.640
RD	24	0.001	0.450
RD	25	0.001	-14.41
RD	26	0.000	-0.880
RD	27	-0.001	-7.698
RD	28	-0.001	0.552
RD	29	-0.001	-0.378
RD	30	0.003	4.715
RD	31	0.001	0.303
RD	32	0.003	5.136
RD	33	0.001	-0.823

Table 13B. Reading Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	34	-0.001	-4.973
RD	35	0.000	2.738

Figure B25. 'A' Anchor Parameters For Reading Items CELApro08 vs. CELApro09/10

Figure B26. 'B' Anchor Parameters For Reading Items CELApro08 vs. CELApro09/10

Table 14B. Speaking Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
SP	1	0.002	1.175
SP	2	0.000	0.037
SP	3	0.000	0.138
SP	4	0.001	0.203
SP	5	0.001	0.461
SP	6	-0.000	-0.191
SP	7	0.004	2.123
SP	8	0.001	0.247
SP	9	0.000	0.258
SP	10	0.000	0.073
SP	11	-0.000	-0.053
SP	12	0.001	0.299
SP	13	0.002	0.810
SP	14	-0.001	-0.481
SP	15	-0.001	-0.308
SP	16	0.000	0.183
SP	17	0.000	0.099
SP	18	0.002	0.878
SP	19	0.000	0.093
SP	20	0.000	0.146

Figure B27. 'A' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10

Figure B28. 'B' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10
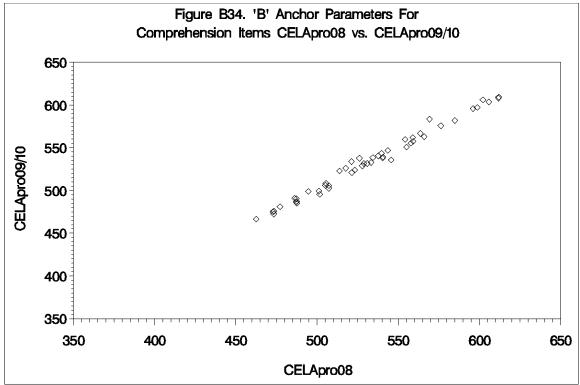
Table 15B. Writing Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
WR	1	0.002	9.941
WR	2	0.001	1.870
WR	3	0.001	5.832
WR	4	0.001	5.278
WR	5	0.002	14.879
WR	6	0.000	2.525
WR	7	-0.001	-0.323
WR	8	0.000	6.482
WR	9	0.001	1.663
WR	10	0.003	1.191
WR	11	0.001	6.204
WR	12	0.004	7.687
WR	13	0.002	6.808
WR	14	0.000	3.435
WR	15	-0.007	-4.872
WR	16	0.004	5.214
WR	17	-0.002	-36.20
WR	18	-0.000	2.765
WR	19	-0.003	-1.031
WR	20	-0.001	4.316

Figure B29. 'A' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Figure B30. 'B' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10
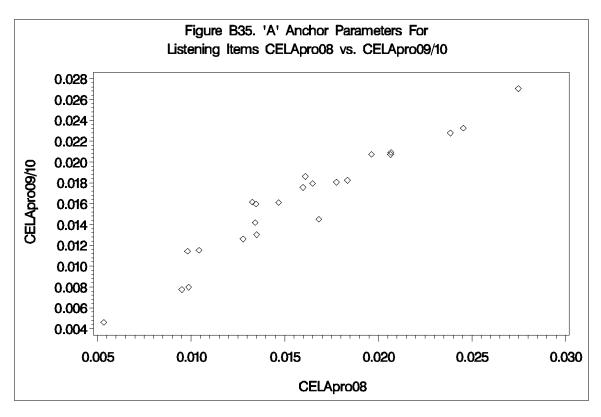
Table 16B. Writing Grade Span 3: 6–8 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
WR	21	0.003	1.108	1.074	1.447	
WR	22	-0.001	-0.448	-0.937	-0.662	
WR	23	0.003	1.139	1.153	1.221	
WR	24	0.001	0.216	0.108	0.173	
WR	25	-0.004	-2.274	-1.427	-2.389	-2.698

Figure B31. 'A' Anchor Parameters For Writing CR Items CELApro08 vs. CELApro09/10

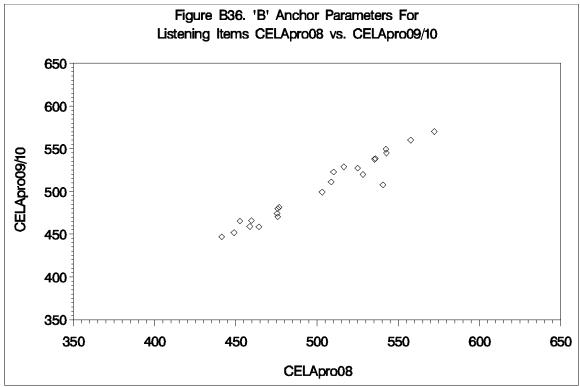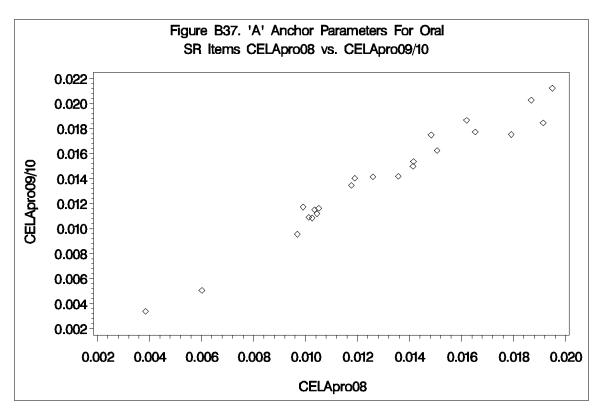Figure B32. 'Gamma' Anchor Parameters For Writing CR Items CELApro08 vs. CELApro09/10

Table 17B. Comprehension Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10

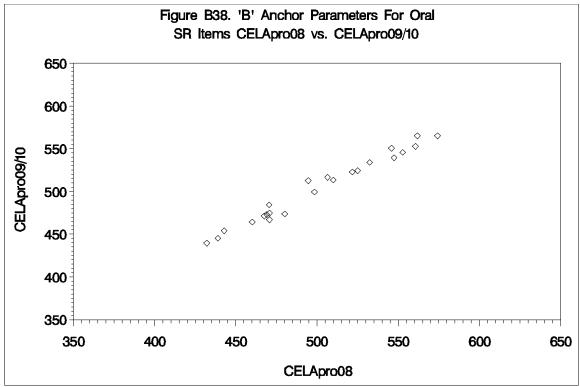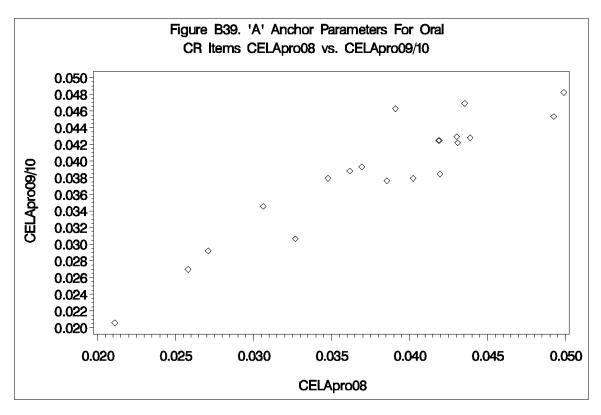| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	1	0.001	2.557
CO	2	-0.001	-1.052
CO	4	0.001	4.022
CO	5	-0.000	3.821
CO	6	0.000	4.180
CO	7	0.000	-0.699
CO	8	-0.000	0.666
CO	10	0.003	5.557
CO	11	0.001	4.970
CO	12	0.000	0.902
CO	13	0.003	9.003
CO	14	-0.001	-3.303
CO	15	-0.000	3.249
CO	17	-0.002	12.505
CO	18	0.002	4.099
CO	19	-0.001	-6.127
CO	20	0.000	1.831
CO	21	-0.001	-3.020
CO	22	-0.000	-1.773
CO	23	-0.001	-10.07
CO	3	0.000	-0.476
CO	5	-0.001	-1.962
CO	6	0.002	-0.601
CO	7	0.001	-0.351
CO	8	0.001	-4.743
CO	9	0.000	-2.352
CO	11	0.000	14.163
CO	12	0.004	8.369
CO	13	-0.001	2.883
CO	14	-0.000	2.583
CO	15	-0.001	-2.793
CO	16	0.001	0.962
CO	18	0.003	-1.601
CO	19	0.001	3.180

Table 17B. Comprehension Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	20	0.000	4.048
CO	21	0.000	3.967
CO	22	-0.001	-2.739
CO	23	0.002	-2.294
CO	24	-0.001	-1.743
CO	25	0.001	2.872
CO	26	0.001	-1.673
CO	27	0.000	-4.369
CO	28	0.003	2.789
CO	29	-0.000	-0.801
CO	30	0.003	-3.616
CO	31	0.003	-0.375
CO	32	-0.000	0.730
CO	33	0.004	11.753
CO	34	0.001	2.726
CO	35	-0.001	-2.503

Figure B33. 'A' Anchor Parameters For
Comprehension Items CELApro08 vs. CELApro09/10

Figure B34. 'B' Anchor Parameters For
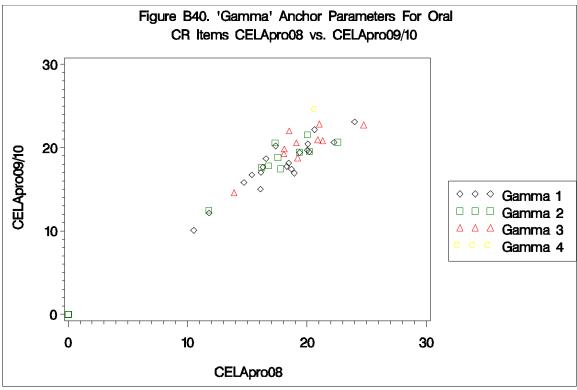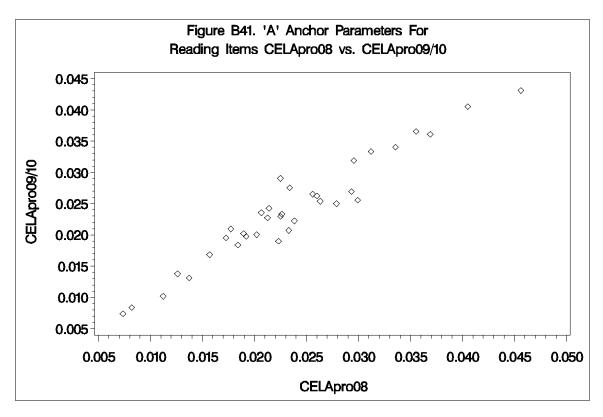Comprehension Items CELApro08 vs. CELApro09/10

Table 18B. Listening Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
LI	1	0.000	-5.663
LI	2	0.000	2.939
LI	3	-0.001	2.461
LI	4	0.001	5.562
LI	5	-0.000	0.425
LI	6	0.000	4.855
LI	7	-0.001	-5.475
LI	8	-0.002	-8.471
LI	9	0.002	2.148
LI	10	0.003	7.075
LI	11	0.001	6.454
LI	12	0.001	4.101
LI	13	0.001	2.502
LI	14	-0.002	-32.72
LI	15	-0.000	2.328
LI	16	0.001	2.598
LI	17	-0.002	12.692
LI	18	-0.000	2.410
LI	19	-0.000	-1.010
LI	20	0.002	12.807
LI	21	0.003	-2.088
LI	22	-0.001	-3.933
LI	23	0.002	12.107

Figure B35. 'A' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Figure B36. 'B' Anchor Parameters For
Listening Items CELApro08 vs. CELApro09/10

Table 19B. Oral Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
OR	1	0.002	4.269
OR	2	0.001	6.319
OR	3	-0.000	3.169
OR	4	0.001	7.294
OR	5	-0.000	-3.994
OR	6	0.001	3.782
OR	7	0.002	4.006
OR	8	0.002	9.840
OR	9	0.001	-7.059
OR	10	0.001	-8.010
OR	11	0.001	-6.299
OR	12	0.002	13.912
OR	13	0.003	3.356
OR	14	-0.000	-7.647
OR	15	0.001	1.451
OR	16	0.002	4.579
OR	17	-0.001	17.884
OR	18	-0.001	1.141
OR	19	0.001	3.790
OR	20	0.002	10.992
OR	21	0.001	-9.003
OR	22	0.002	0.984
OR	23	0.001	-0.854

Figure B37. 'A' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10

Figure B38. 'B' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10
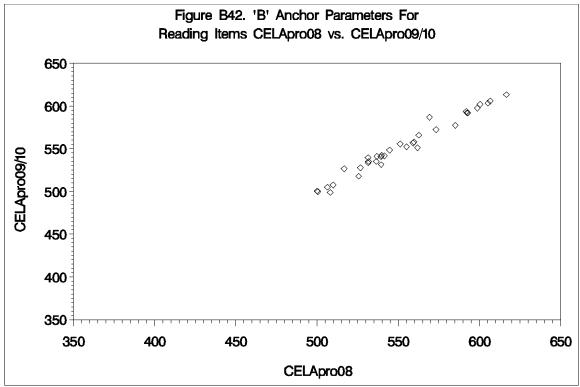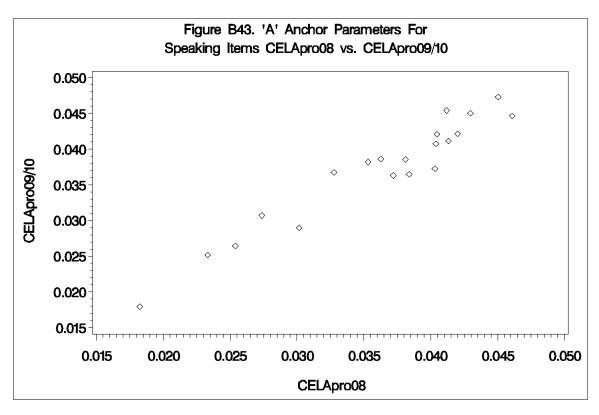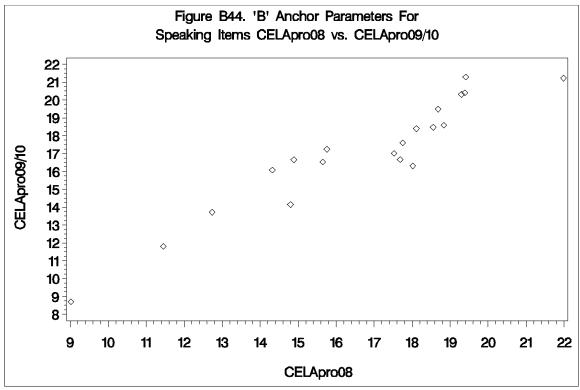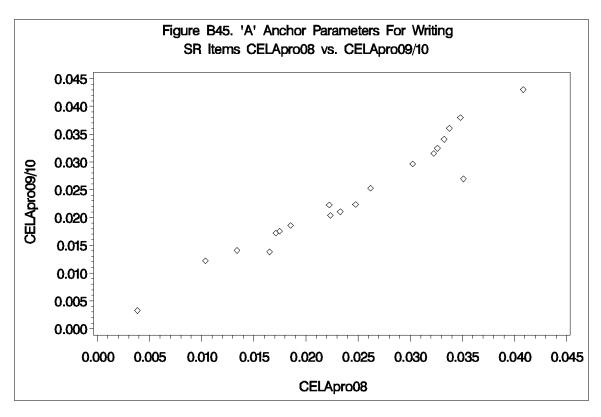
Table 20B. Oral Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
OR	1	0.001	0.360	0.000		
OR	2	-0.001	-0.704	0.000		
OR	3	0.004	2.115	0.000		
OR	4	-0.002	-1.253	0.000		
OR	5	0.002	1.130	0.000		
OR	6	-0.002	-1.077	0.000		
OR	7	-0.002	-0.883	0.000		
OR	8	0.001	0.016	0.000		
OR	9	-0.001	-0.434	0.000		
OR	10	-0.004	-1.974	0.000		
OR	11	0.002	0.897	1.043	1.180	
OR	12	-0.000	-0.328	0.081	0.051	
OR	13	0.001	0.361	0.727	0.696	
OR	14	0.003	1.361	1.274	1.471	
OR	15	0.003	1.350	1.432	1.760	
OR	16	0.003	1.516	1.519	1.814	
OR	17	-0.001	-0.562	-0.321	-0.454	
OR	18	-0.001	-0.341	-0.662	-0.464	
OR	19	-0.004	-1.627	-1.905	-2.019	
OR	20	0.007	2.843	3.241	3.519	3.877

Figure B39. 'A' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

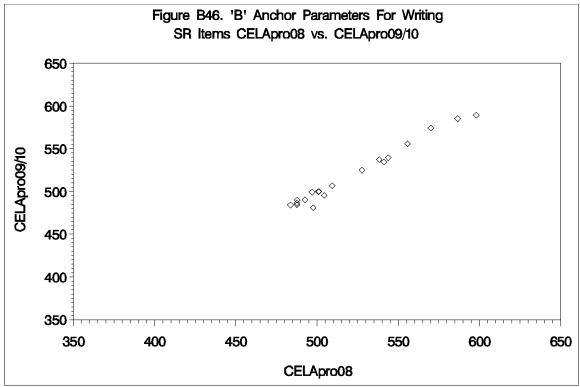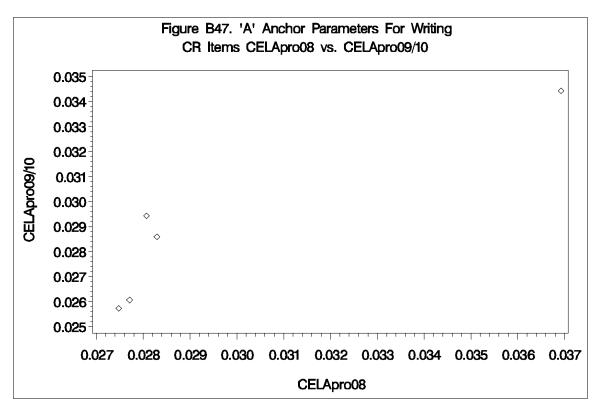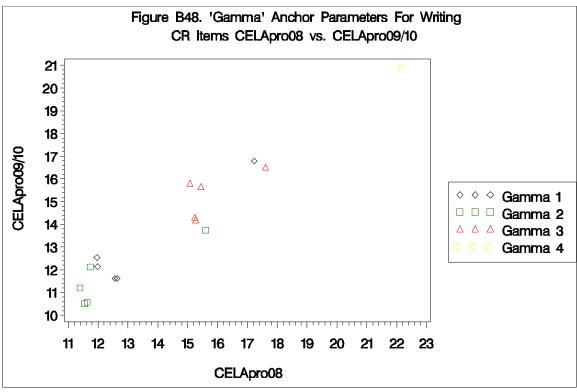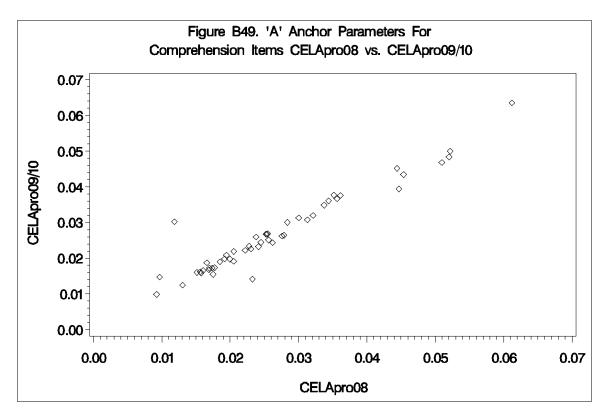Figure B40. 'Gamma' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

Table 21B. Reading Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	1	-0.001	-1.060
RD	2	0.000	-7.848
RD	3	0.000	0.684
RD	4	0.007	-3.044
RD	5	-0.003	-2.279
RD	6	0.001	-1.020
RD	7	0.000	0.780
RD	8	0.001	-1.401
RD	9	0.001	1.337
RD	10	-0.001	-10.76
RD	11	0.001	17.691
RD	12	0.004	9.779
RD	13	-0.003	-9.205
RD	14	-0.004	-7.693
RD	15	-0.001	-0.655
RD	16	0.001	1.959
RD	17	-0.000	3.653
RD	18	0.003	-0.067
RD	19	0.001	3.397
RD	20	-0.000	4.132
RD	21	0.000	1.587
RD	22	-0.003	-0.671
RD	23	0.002	-1.307
RD	24	-0.002	-1.857
RD	25	0.001	4.313
RD	26	0.002	0.330
RD	27	0.000	-2.861
RD	28	0.003	3.467
RD	29	-0.001	-1.079
RD	30	0.002	-1.990
RD	31	0.001	1.673
RD	32	-0.003	-8.078
RD	33	0.003	8.429

Table 21B. Reading Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	34	0.000	2.384
RD	35	-0.002	-2.272

Figure B41. 'A' Anchor Parameters For
Reading Items CELApro08 vs. CELApro09/10

Figure B42. 'B' Anchor Parameters For
Reading Items CELApro08 vs. CELApro09/10

Table 22B. Speaking Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
SP	1	0.000	0.284
SP	2	-0.000	-0.247
SP	3	0.003	1.768
SP	4	-0.002	-1.021
SP	5	0.002	0.978
SP	6	-0.001	-0.660
SP	7	-0.001	-0.762
SP	8	0.000	-0.082
SP	9	-0.000	-0.315
SP	10	-0.003	-1.711
SP	11	0.002	0.893
SP	12	0.000	-0.154
SP	13	0.001	0.358
SP	14	0.003	1.484
SP	15	0.004	1.761
SP	16	0.004	1.880
SP	17	-0.001	-0.508
SP	18	0.002	0.808
SP	19	0.002	1.027
SP	20	0.002	1.032

Figure B43. 'A' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10

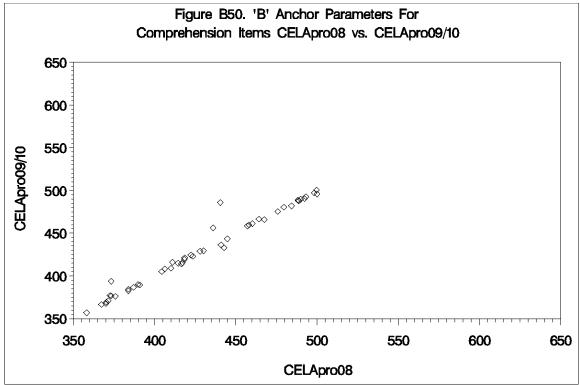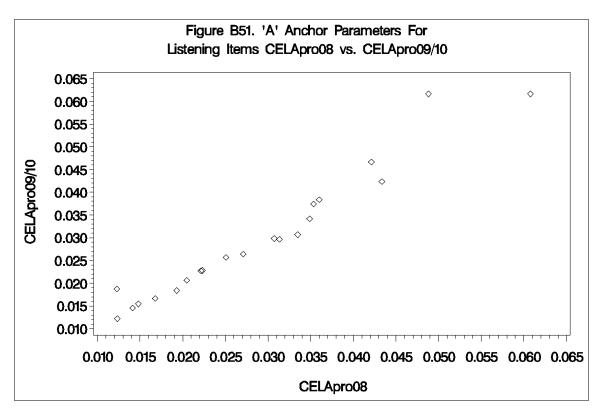Figure B44. 'B' Anchor Parameters For Speaking Items CELApro08 vs. CELApro09/10

Table 23B. Writing Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
WR	1	0.000	0.235
WR	2	-0.001	-2.428
WR	3	0.002	2.199
WR	4	0.003	0.408
WR	5	0.001	-1.363
WR	6	-0.000	-2.407
WR	7	0.000	-0.878
WR	8	0.002	2.183
WR	9	-0.001	-1.125
WR	10	-0.001	5.982
WR	11	-0.001	-3.275
WR	12	-0.002	-6.437
WR	13	-0.008	-9.036
WR	14	-0.002	-1.102
WR	15	0.000	4.445
WR	16	-0.002	-2.821
WR	17	-0.003	-16.600
WR	18	0.001	-1.364
WR	19	0.002	-8.569
WR	20	0.000	-4.585

Figure B45. 'A' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Figure B46. 'B' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

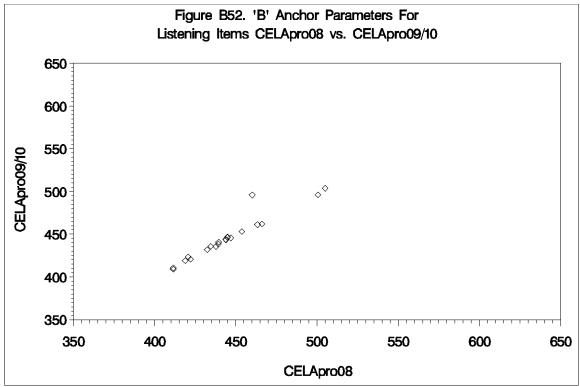Table 24B. Writing Grade Span 4: 9–12 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
WR	21	0.000	0.156	-0.190	0.223	
WR	22	0.001	0.569	0.391	0.734	
WR	23	-0.002	-1.009	-1.054	-1.068	
WR	24	-0.002	-0.941	-1.038	-0.918	
WR	25	-0.002	-0.441	-1.854	-1.083	-1.274

Figure B47. 'A' Anchor Parameters For Writing
CR Items CELApro08 vs. CELApro09/10

Figure B48. 'Gamma' Anchor Parameters For Writing
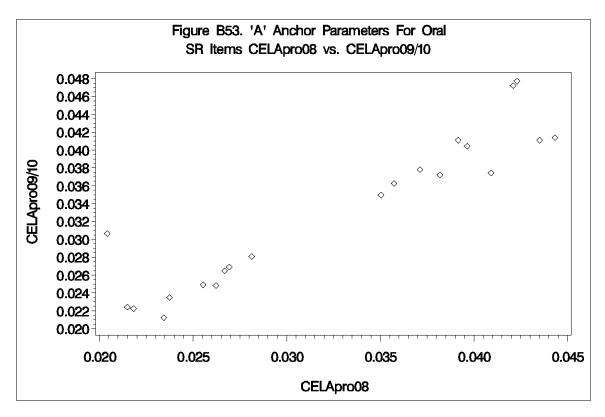CR Items CELApro08 vs. CELApro09/10

Table 25B. Comprehension Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10

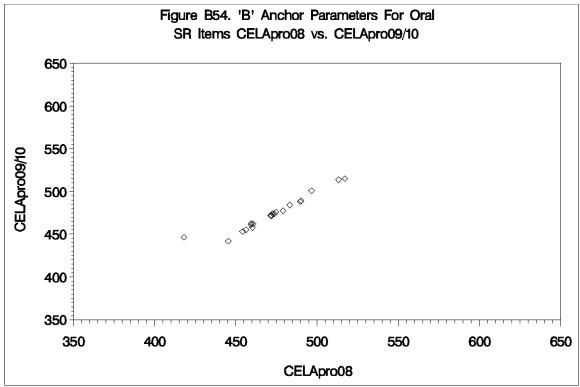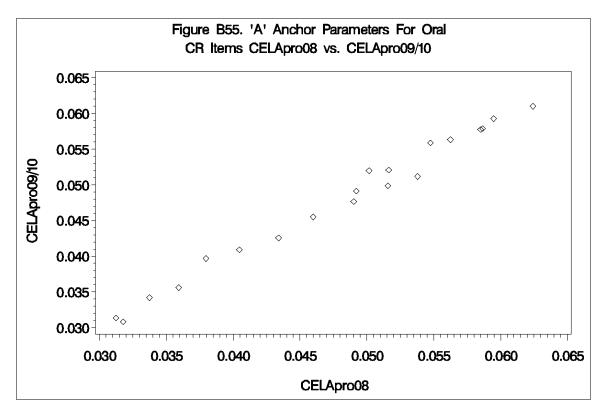| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	1	-0.002	-2.294
CO	2	-0.000	-1.280
CO	3	-0.000	-1.243
CO	4	-0.000	-0.623
CO	5	0.002	-0.212
CO	6	0.000	0.700
CO	7	0.001	0.577
CO	8	0.003	0.916
CO	9	-0.000	0.330
CO	11	0.005	45.250
CO	12	0.000	2.031
CO	13	-0.001	-0.628
CO	14	-0.001	-4.603
CO	15	-0.000	-1.450
CO	16	0.001	-1.020
CO	17	0.001	-1.708
CO	18	0.001	2.410
CO	19	0.000	1.953
CO	20	0.001	1.235
CO	21	0.000	0.495
CO	1	-0.001	-2.226
CO	2	-0.001	-3.438
CO	3	-0.001	-1.413
CO	4	-0.000	-0.397
CO	8	0.002	2.361
CO	9	0.000	0.788
CO	10	0.001	4.968
CO	11	-0.000	-0.715
CO	20	-0.004	-0.394
CO	21	-0.005	0.655
CO	22	0.002	20.342
CO	23	-0.000	0.628
CO	24	0.001	-0.832
CO	25	-0.001	-1.720

Table 25B. Comprehension Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
CO	26	0.001	0.985
CO	27	-0.009	-9.559
CO	27	0.001	1.752
CO	28	0.002	4.632
CO	28	0.001	-0.460
CO	29	0.001	3.411
CO	29	-0.002	-1.544
CO	30	0.018	19.978
CO	30	0.001	1.090
CO	31	0.001	-1.107
CO	31	0.002	-0.874
CO	32	0.002	0.739
CO	33	-0.002	-0.480
CO	34	0.001	-0.521
CO	35	-0.002	-4.380
CO	36	-0.004	-2.257

Figure B49. 'A' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10

Figure B50. 'B' Anchor Parameters For Comprehension Items CELApro08 vs. CELApro09/10
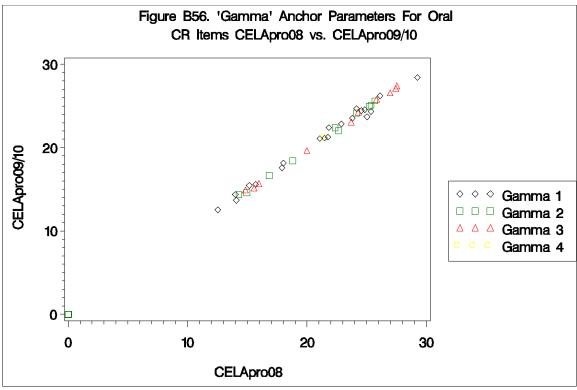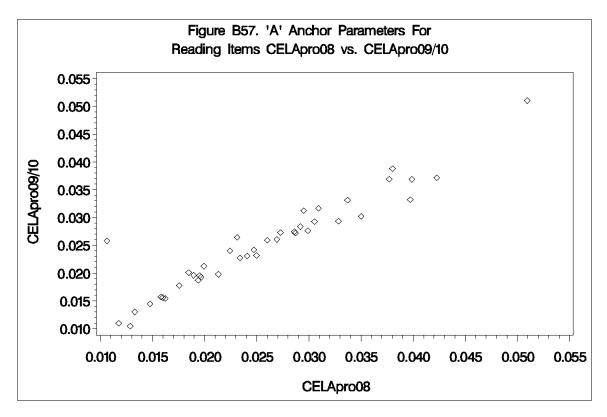
Table 26B. Listening Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
LI	1	-0.003	-2.268
LI	2	-0.001	-1.338
LI	3	-0.001	-0.938
LI	4	0.005	1.534
LI	5	0.002	0.275
LI	6	0.001	-0.351
LI	7	0.001	-0.700
LI	8	0.013	1.743
LI	9	0.000	2.584
LI	10	0.001	-0.743
LI	11	0.006	35.784
LI	12	-0.001	0.128
LI	13	-0.002	-1.318
LI	14	-0.001	-3.977
LI	15	-0.000	-1.030
LI	16	0.002	-1.195
LI	17	-0.000	-4.436
LI	18	-0.001	-2.092
LI	19	0.001	1.212
LI	20	0.000	1.577
LI	21	0.000	-2.373

Figure B51. 'A' Anchor Parameters For Listening Items CELApro08 vs. CELApro09/10

Figure B52. 'B' Anchor Parameters For Listening Items CELApro08 vs. CELApro09/10

Table 27B. Oral Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
OR	1	-0.003	0.250
OR	2	-0.003	-2.431
OR	3	-0.002	-3.322
OR	4	0.010	28.316
OR	5	0.001	-1.068
OR	6	-0.000	-0.397
OR	7	0.005	2.003
OR	8	0.005	2.963
OR	9	0.001	1.163
OR	10	-0.000	-1.048
OR	11	-0.002	4.322
OR	12	0.001	1.329
OR	13	-0.001	-0.063
OR	14	-0.001	-1.217
OR	15	-0.001	-1.753
OR	16	-0.000	0.529
OR	17	0.001	-2.125
OR	18	-0.000	-0.034
OR	19	0.002	1.859
OR	20	0.000	0.938
OR	21	-0.000	-1.244

Figure B53. 'A' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10

Figure B54. 'B' Anchor Parameters For Oral SR Items CELApro08 vs. CELApro09/10
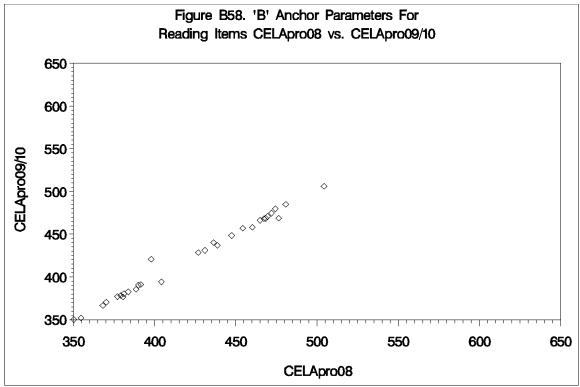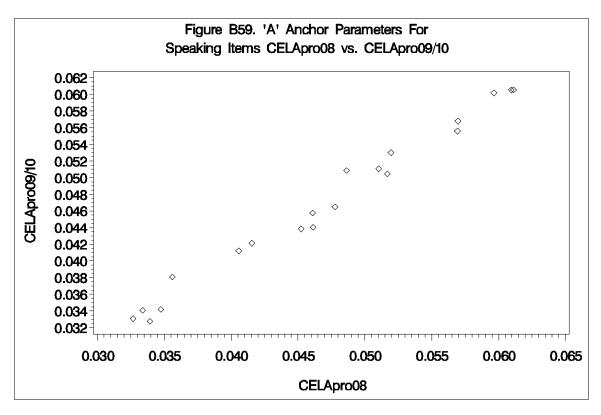
Table 28B. Oral Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
OR	1	0.002	0.381	0.000		
OR	2	0.000	0.121	0.000		
OR	3	0.001	0.531	0.000		
OR	4	0.002	0.572	0.000		
OR	5	0.000	0.306	0.000		
OR	6	-0.001	-0.311	0.000		
OR	7	-0.003	-1.313	0.000		
OR	8	-0.001	-0.848	0.000		
OR	9	0.000	0.096	0.000		
OR	10	-0.002	-1.035	0.000		
OR	11	-0.000	-0.110	-0.044	-0.121	
OR	12	-0.001	-0.273	-0.272	-0.316	
OR	13	0.000	-0.037	0.017	-0.008	
OR	14	-0.001	-0.295	-0.266	-0.373	
OR	15	-0.000	0.022	0.014	-0.012	
OR	16	0.000	0.021	0.078	0.060	
OR	17	-0.000	-0.108	-0.177	-0.270	
OR	18	-0.001	-0.393	-0.369	-0.410	
OR	19	-0.001	-0.496	-0.595	-0.614	
OR	20	-0.001	-0.335	-0.365	-0.322	-0.304

Figure B55. 'A' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10

Figure B56. 'Gamma' Anchor Parameters For Oral CR Items CELApro08 vs. CELApro09/10
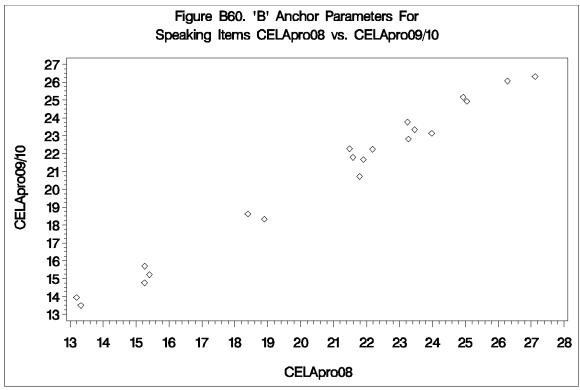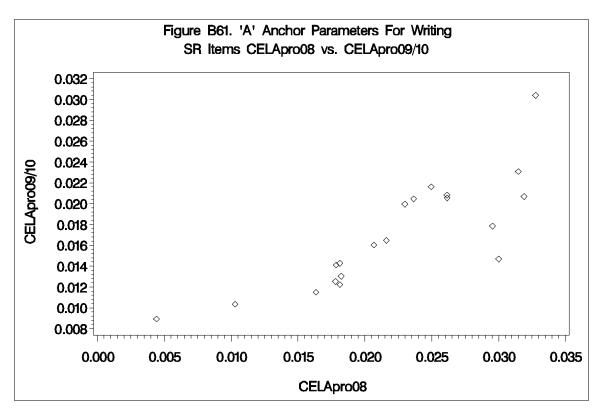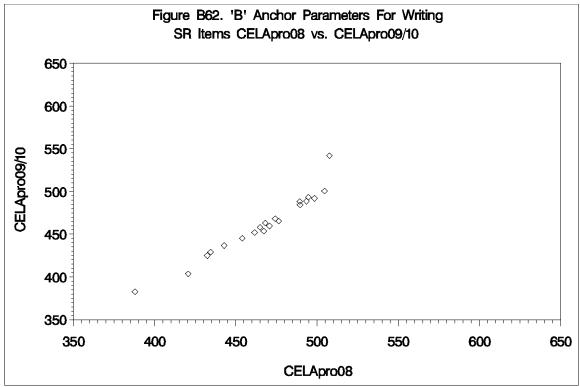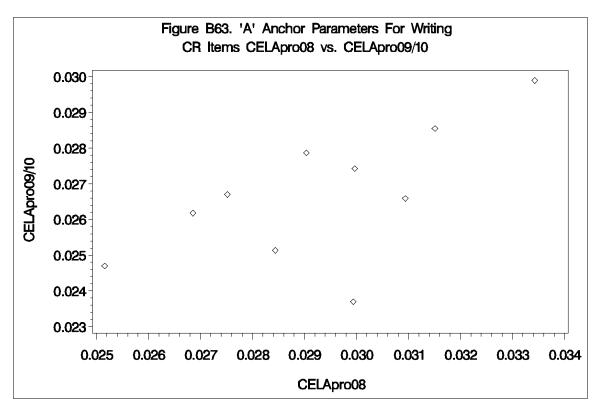
Table 29B. Reading Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10

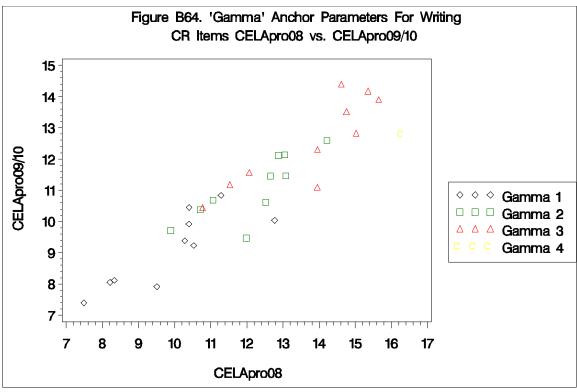| Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	1	-0.001	-1.878
RD	2	-0.002	-5.240
RD	3	-0.001	-0.234
RD	4	-0.001	-0.985
RD	5	-0.001	-0.792
RD	6	0.003	5.426
RD	7	0.001	2.022
RD	8	0.001	3.799
RD	9	-0.000	0.115
RD	10	-0.000	0.226
RD	11	-0.001	-0.392
RD	12	0.000	-0.206
RD	13	0.000	-0.665
RD	14	0.002	-0.798
RD	15	-0.002	-3.059
RD	16	0.001	-0.195
RD	17	-0.001	-1.176
RD	18	-0.001	-0.884
RD	19	-0.001	-2.680
RD	20	-0.005	1.308
RD	21	-0.003	4.242
RD	22	-0.000	-2.388
RD	23	-0.000	0.103
RD	24	0.000	-1.563
RD	25	-0.002	-3.862
RD	26	-0.000	0.040
RD	27	-0.001	-9.953
RD	27	-0.001	1.681
RD	28	0.002	5.638
RD	28	-0.000	1.005
RD	29	0.001	4.497
RD	29	-0.003	-1.602
RD	30	0.015	22.928
RD	30	-0.001	2.633

Table 29B. Reading Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (continued)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
RD	31	0.002	-1.448
RD	31	-0.001	0.258
RD	32	0.000	1.392
RD	33	-0.005	0.876
RD	34	-0.001	2.562
RD	35	-0.002	-7.673
RD	36	-0.006	-2.042

Figure B57. 'A' Anchor Parameters For
Reading Items CELApro08 vs. CELApro09/10

Figure B58. 'B' Anchor Parameters For
Reading Items CELApro08 vs. CELApro09/10

Table 30B. Speaking Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
SP	1	0.002	0.754
SP	2	0.001	0.218
SP	3	0.001	0.522
SP	4	0.002	0.776
SP	5	0.001	0.430
SP	6	-0.000	-0.235
SP	7	-0.002	-1.068
SP	8	-0.001	-0.808
SP	9	0.001	0.204
SP	10	-0.001	-0.839
SP	11	0.001	0.217
SP	12	-0.000	-0.113
SP	13	-0.000	-0.130
SP	14	-0.001	-0.220
SP	15	0.000	0.059
SP	16	0.000	0.176
SP	17	-0.001	-0.199
SP	18	-0.001	-0.499
SP	19	-0.001	-0.452
SP	20	-0.001	-0.566

Figure B59. 'A' Anchor Parameters For
Speaking Items CELApro08 vs. CELApro09/10

Figure B60. 'B' Anchor Parameters For
Speaking Items CELApro08 vs. CELApro09/10

Table 31B. Writing Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (SR Items)

Content	Item Number	"A" Parameter Difference	"B" Parameter Difference
WR	1	-0.003	-7.544
WR	2	-0.003	-8.715
WR	3	-0.002	-5.303
WR	4	-0.008	-6.147
WR	5	-0.005	-9.515
WR	6	-0.006	-17.130
WR	7	0.000	-5.179
WR	8	-0.005	-11.120
WR	9	-0.015	-4.891
WR	10	-0.004	-13.380
WR	11	-0.003	-6.109
WR	12	0.005	34.118
WR	13	-0.005	-6.557
WR	14	-0.005	-5.027
WR	15	-0.006	-11.190
WR	16	-0.012	-1.198
WR	17	-0.004	-1.349
WR	18	-0.005	-3.972
WR	19	-0.011	-5.202
WR	20	-0.005	-6.641

Figure B61. 'A' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Figure B62. 'B' Anchor Parameters For Writing
SR Items CELApro08 vs. CELApro09/10

Table 32B. Writing Grade Span 1: K–2 CELApro 08 vs. CELApro 09/10 (CR Items)

Content	Item Number	"A" Parameter Difference	"B/Gamma 1" Parameter Difference	"Gamma 2" Parameter Difference	"Gamma 3" Parameter Difference	"Gamma 4" Parameter Difference
WR	21	-0.004	-1.595	-1.908	-2.197	
WR	21	-0.003	-0.447	-1.626	-1.177	
WR	22	-0.004	-1.297	-1.612	-1.748	
WR	22	-0.001	0.047	-0.761	-0.218	
WR	23	-0.000	-0.09/103	-0.181	-0.327	
WR	23	-0.003	-0.474	-0.916	-1.238	
WR	24	-0.001	-0.216	-0.382	-0.494	
WR	24	-0.003	-0.898	-1.201	-1.649	
WR	25	-0.001	-0.154	-0.331	-0.346	
WR	25	-0.006	-2.730	-2.515	-2.851	-3.489

Figure B63. 'A' Anchor Parameters For Writing
CR Items CELApro08 vs. CELApro09/10

Figure B64. 'Gamma' Anchor Parameters For Writing
CR Items CELApro08 vs. CELApro09/10

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix C: TCC & SEM Plots by Grade Span

List of Figures

Figure 1C CELApro 09/10 Speaking TCC Plots by Grade Span

Speaking: Level 1 = Grades K–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 2C CELApro 09/10 Speaking SEM Plots by Grade Span

Speaking: Level 1 = Grades K–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 3C CELApro 09/10 Listening TCC Plots by Grade Span

Listening: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 4C CELApro 09/10 Listening SEM Plots by Grade Span

Listening: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 5C CELApro 09/10 Reading TCC Plots by Grade Span

Reading: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 6C CELApro 09/10 Reading SEM Plots by Grade Span

Reading: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 7C CELApro 09/10 Writing TCC Plots by Grade Span

Writing: Level 0 = Grades K–1, Level 1 = Grade 2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 8C CELApro 09/10 Writing SEM Plots by Grade Span

Writing: Level 0 = Grades K–1, Level 1 = Grade 2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 9C CELApro 09/10 Comprehension TCC Plots by Grade Span

Comprehension: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 10C CELApro 09/10 Comprehension SEM Plots by Grade Span

Comprehension: Level 0 = Grade K, Level 1 = Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 11C CELApro 09/10 Oral TCC Plots by Grade Span

Oral: Level 0 = Grade K, Level 1 = Grades K–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Figure 12C CELApro 09/10 Oral SEM Plots by Grade Span

Oral: Level 0 = Grade K, Level 1= Grades 1–2, Level 2 = Grades 3–5, Level 3 = Grades 6–8, and Level 4 = Grades 9–12

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix D: Equating Results for Grade Spans

List of Figures

Figure 1D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Speaking

Figure 2D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Listening

Figure 3D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Reading

Figure 4D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Writing

Figure 5D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Comprehension

Figure 6D. Stocking & Lord Test Characteristic Curves, Grade Span 1: K–2 Oral

Figure 7D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Speaking

Figure 8D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Listening

Figure 9D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Reading

Figure 10D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Writing

Figure 11D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Comprehension

Figure 12D. Stocking & Lord Test Characteristic Curves, Grade Span 2: 3–5 Oral

Figure 13D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Speaking

Figure 14D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Listening

Figure 15D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Reading

Figure 16D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Writing

Figure 17D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Comprehension

Figure 18D. Stocking & Lord Test Characteristic Curves, Grade Span 3: 6–8 Oral

Figure 19D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Speaking

Figure 20D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Listening

Figure 21D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Reading

Figure 22D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Writing

Figure 23D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Comprehension

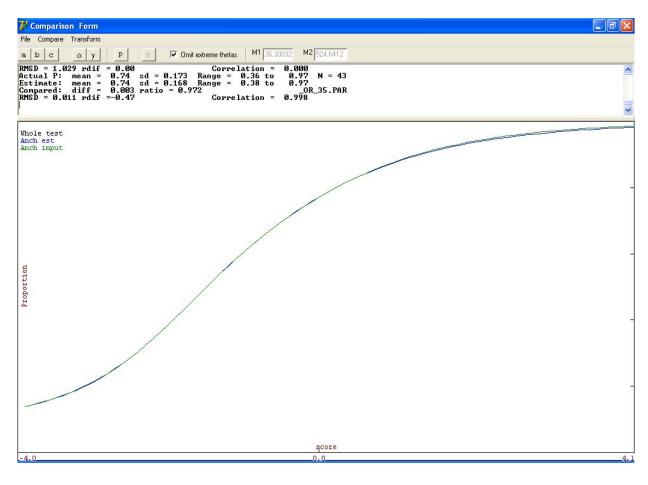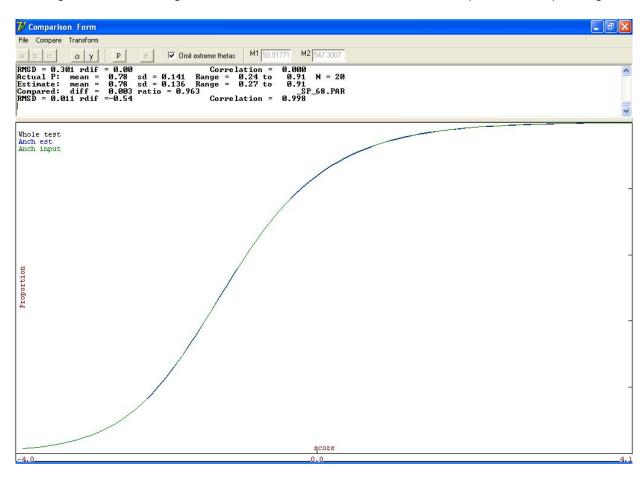Figure 24D. Stocking & Lord Test Characteristic Curves, Grade Span 4: 9–12 Oral

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix E: Raw Score to Scale Score Tables

List of Tables

Table 1E. Raw Score to Scale Score Table Grade Span 1: K, Speaking

Raw Score	Scale Score	Standard Error
0	300	76
1	359	27
2	380	18
3	391	14
4	400	12
5	406	11
6	411	10
7	416	9
8	420	9
9	424	9
10	427	8
11	430	8
12	433	8
13	436	8
14	439	8
15	442	8
16	445	8
17	447	8
18	450	8
19	453	8
20	455	8
21	458	8
22	461	8
23	463	8
24	466	8
25	469	8
26	471	8
27	474	8
28	477	8
29	480	8
30	483	9
31	487	9
32	490	9
33	494	10
34	499	10
35	504	11
36	509	12
37	516	13
38	524	15
39	535	18
40	553	25
41	580	43

Table 2E. Raw Score to Scale Score Table Grade Span 1: K, Listening

Raw Score	Scale Score	Standard Error
0	300	111
1	300	111
2	300	111
3	327	84
4	373	39
5	390	24
6	401	19
7	410	16
8	417	15
9	424	14
10	430	13
11	436	13
12	442	12
13	448	13
14	455	13
15	463	15
16	472	17
17	484	21
18	502	27
19	533	42
20	560	58

Table 3E. Raw Score to Scale Score Table Grade Span 1: K, Reading

Raw Score	Scale Score	Standard Error
0	240	77
1	240	77
2	240	77
3	240	77
4	240	77
5	240	77
6	240	77
7	270	47
8	287	33
9	299	26
10	309	23
11	318	21
12	326	19
13	333	18
14	340	17
15	347	16
16	353	16
17	359	15
18	365	15
19	371	14
20	377	14
21	383	15
22	390	15
23	397	16
24	405	17
25	414	19
26	425	22
27	439	25
28	457	27
29	478	31
30	523	71
31	570	118

Table 4E. Raw Score to Scale Score Table Grade Span 1: K, Writing

Raw Score	Scale Score	Standard Error
0	200	139
1	200	139
2	200	139
3	200	139
4	200	139
5	200	139
6	210	129
7	264	74
8	292	53
9	312	45
10	330	41
11	346	37
12	360	34
13	373	30
14	384	27
15	394	24
16	403	22
17	411	21
18	418	19
19	425	18
20	432	18
21	439	17
22	445	17
23	452	16
24	458	16
25	464	16
26	471	16
27	478	16
28	485	16
29	492	16
30	500	17
31	510	19
32	522	21
33	539	27
34	569	44
35	630	105

Table 5E. Raw Score to Scale Score Table Grade Span 1: K, Comprehension

Raw Score	Scale Score	Standard Error
0	270	85
1	270	85
2	270	85
3	270	85
4	270	85
5	270	85
6	270	85
7	293	62
8	316	39
9	329	28
10	339	23
11	348	20
12	355	18
13	361	17
14	367	16
15	372	15
16	377	14
17	382	14
18	387	14
19	392	13
20	396	13
21	401	13
22	406	13
23	410	13
24	415	13
25	420	13
26	425	13
27	430	14
28	435	14
29	441	15
30	447	15
31	454	16
32	462	17
33	470	18
34	480	18
35	490	19
36	503	22
37	521	29
38	553	43
39	570	54

Table 6E. Raw Score to Scale Score Table Grade Span 1: K, Oral

Raw Score	Scale Score	Standard Error
0	280	125
1	280	125
2	280	125
3	280	125
4	280	125
5	280	125
6	280	125
7	345	60
8	368	37
9	381	25
10	389	19
11	396	16
12	401	14
13	406	13
14	410	11
15	413	11
16	417	10
17	420	9
18	422	9
19	425	9
20	427	8
21	430	8
22	432	8
23	434	8
24	437	8
25	439	7
26	441	7
27	443	7
28	445	7
29	447	7
30	449	7
31	451	7
32	452	7
33	454	7
34	456	7
35	458	7
36	460	7
37	462	7
38	464	7
39	465	7
40	467	7
41	469	7
42	471	7
43	473	7
44	476	7
45	478	7

Table 6E. Raw Score to Scale Score Table Grade Span 1: K, Oral (continued)

Raw Score	Scale Score	Standard Error
46	480	7
47	482	7
48	485	8
49	487	8
50	490	8
51	493	8
52	496	9
53	500	9
54	504	10
55	508	10
56	514	11
57	520	13
58	527	15
59	538	18
60	556	25
61	620	89

Table 7E. Raw Score to Scale Score Table Grade Span 1: 1, Speaking

Raw Score	Scale Score	Standard Error
0	300	76
1	359	27
2	380	18
3	391	14
4	400	12
5	406	11
6	411	10
7	416	9
8	420	9
9	424	9
10	427	8
11	430	8
12	433	8
13	436	8
14	439	8
15	442	8
16	445	8
17	447	8
18	450	8
19	453	8
20	455	8
21	458	8
22	461	8
23	463	8
24	466	8
25	469	8
26	471	8
27	474	8
28	477	8
29	480	8
30	483	9
31	487	9
32	490	9
33	494	10
34	499	10
35	504	11
36	509	12
37	516	13
38	524	15
39	535	18
40	553	25
41	580	43

Table 8E. Raw Score to Scale Score Table Grade Span 1: 1, Listening

Raw Score	Scale Score	Standard Error
0	300	113
1	300	113
2	300	113
3	300	113
4	361	52
5	384	29
6	397	21
7	406	18
8	414	16
9	421	14
10	427	14
11	434	13
12	439	13
13	446	13
14	452	13
15	459	14
16	467	16
17	478	19
18	491	22
19	508	26
20	537	38
21	560	54

Table 9E. Raw Score to Scale Score Table Grade Span 1: 1, Reading

Raw Score	Scale Score	Standard Error
0	240	96
1	240	96
2	240	96
3	240	96
4	240	96
5	240	96
6	240	96
7	240	96
8	251	85
9	282	54
10	300	38
11	313	31
12	325	27
13	334	24
14	343	22
15	352	21
16	359	20
17	367	19
18	374	18
19	381	18
20	388	18
21	396	19
22	404	19
23	412	20
24	421	19
25	429	18
26	437	17
27	445	16
28	452	15
29	459	14
30	466	13
31	473	13
32	481	14
33	491	16
34	503	19
35	523	27
36	590	94

Table 10E. Raw Score to Scale Score Table Grade Span 1: 1, Writing

Raw Score	Scale Score	Standard Error
0	200	139
1	200	139
2	200	139
3	200	139
4	200	139
5	200	139
6	210	129
7	264	74
8	292	53
9	312	45
10	330	41
11	346	37
12	360	34
13	373	30
14	384	27
15	394	24
16	403	22
17	411	21
18	418	19
19	425	18
20	432	18
21	439	17
22	445	17
23	452	16
24	458	16
25	464	16
26	471	16
27	478	16
28	485	16
29	492	16
30	500	17
31	510	19
32	522	21
33	539	27
34	569	44
35	630	105

Table 11E. Raw Score to Scale Score Table Grade Span 1: 1, Comprehension

Raw Score	Scale Score	Standard Error
0	270	97
1	270	97
2	270	97
3	270	97
4	270	97
5	270	97
6	270	97
7	270	97
8	270	97
9	282	85
10	313	53
11	330	37
12	341	29
13	351	25
14	359	22
15	366	20
16	372	19
17	378	18
18	384	17
19	390	17
20	395	16
21	400	16
22	406	16
23	411	15
24	416	15
25	421	15
26	427	15
27	432	15
28	437	15
29	443	15
30	449	15
31	454	14
32	459	14
33	465	13
34	470	12
35	475	12
36	479	11
37	484	11
38	489	11
39	494	11
40	500	12
41	506	13
42	515	16
43	528	22
44	554	39
45	590	69

Table 12E. Raw Score to Scale Score Table Grade Span 1: 1, Oral

Raw Score	Scale Score	Standard Error
0	280	126
1	280	126
2	280	126
3	280	126
4	280	126
5	280	126
6	280	126
7	329	77
8	362	44
9	377	29
10	386	21
11	394	17
12	399	15
13	404	13
14	408	12
15	412	11
16	415	10
17	418	10
18	421	9
19	424	9
20	427	9
21	429	8
22	431	8
23	434	8
24	436	8
25	438	8
26	440	7
27	442	7
28	444	7
29	446	7
30	448	7
31	450	7
32	452	7
33	454	7
34	455	7
35	457	7
36	459	7
37	461	7
38	463	7
39	465	7
40	467	7
41	469	7
42	470	7
43	472	7
44	475	7
45	477	7

Table 12E. Raw Score to Scale Score Table Grade Span 1: 1, Oral (continued)

Raw Score	Scale Score	Standard Error
46	479	7
47	481	7
48	484	7
49	486	8
50	489	8
51	492	8
52	495	9
53	498	9
54	501	9
55	505	10
56	510	11
57	515	12
58	522	13
59	529	15
60	540	18
61	558	25
62	620	87

Table 13E. Raw Score to Scale Score Table Grade Span 1: 2, Speaking

Raw Score	Scale Score	Standard Error
0	300	76
1	359	27
2	380	18
3	391	14
4	400	12
5	406	11
6	411	10
7	416	9
8	420	9
9	424	9
10	427	8
11	430	8
12	433	8
13	436	8
14	439	8
15	442	8
16	445	8
17	447	8
18	450	8
19	453	8
20	455	8
21	458	8
22	461	8
23	463	8
24	466	8
25	469	8
26	471	8
27	474	8
28	477	8
29	480	8
30	483	9
31	487	9
32	490	9
33	494	10
34	499	10
35	504	11
36	509	12
37	516	13
38	524	15
39	535	18
40	553	25
41	580	43

Table 14E. Raw Score to Scale Score Table Grade Span 1: 2, Listening

Raw Score	Scale Score	Standard Error
0	300	113
1	300	113
2	300	113
3	300	113
4	361	52
5	384	29
6	397	21
7	406	18
8	414	16
9	421	14
10	427	14
11	434	13
12	439	13
13	446	13
14	452	13
15	459	14
16	467	16
17	478	19
18	491	22
19	508	26
20	537	38
21	560	54

Table 15E. Raw Score to Scale Score Table Grade Span 1: 2, Reading

Raw Score	Scale Score	Standard Error
0	240	96
1	240	96
2	240	96
3	240	96
4	240	96
5	240	96
6	240	96
7	240	96
8	251	85
9	282	54
10	300	38
11	313	31
12	325	27
13	334	24
14	343	22
15	352	21
16	359	20
17	367	19
18	374	18
19	381	18
20	388	18
21	396	19
22	404	19
23	412	20
24	421	19
25	429	18
26	437	17
27	445	16
28	452	15
29	459	14
30	466	13
31	473	13
32	481	14
33	491	16
34	503	19
35	523	27
36	590	94

Table 16E. Raw Score to Scale Score Table Grade Span 1: 2, Writing

Raw Score	Scale Score	Standard Error
0	200	189
1	200	189
2	200	189
3	200	189
4	200	189
5	200	189
6	261	128
7	317	72
8	343	49
9	361	39
10	376	32
11	388	27
12	398	24
13	406	21
14	414	19
15	421	18
16	427	17
17	433	16
18	439	16
19	445	15
20	450	15
21	456	15
22	461	14
23	466	14
24	472	14
25	477	14
26	483	14
27	489	14
28	495	15
29	502	15
30	509	16
31	518	17
32	528	19
33	540	23
34	558	29
35	590	45
36	640	94

Table 17E. Raw Score to Scale Score Table Grade Span 1: 2, Comprehension

Raw Score	Scale Score	Standard Error
0	270	97
1	270	97
2	270	97
3	270	97
4	270	97
5	270	97
6	270	97
7	270	97
8	270	97
9	282	85
10	313	53
11	330	37
12	341	29
13	351	25
14	359	22
15	366	20
16	372	19
17	378	18
18	384	17
19	390	17
20	395	16
21	400	16
22	406	16
23	411	15
24	416	15
25	421	15
26	427	15
27	432	15
28	437	15
29	443	15
30	449	15
31	454	14
32	459	14
33	465	13
34	470	12
35	475	12
36	479	11
37	484	11
38	489	11
39	494	11
40	500	12
41	506	13
42	515	16
43	528	22
44	554	39
45	590	69

Table 18E. Raw Score to Scale Score Table Grade Span 1: 2, Oral

Raw Score	Scale Score	Standard Error
0	280	126
1	280	126
2	280	126
3	280	126
4	280	126
5	280	126
6	280	126
7	329	77
8	362	44
9	377	29
10	386	21
11	394	17
12	399	15
13	404	13
14	408	12
15	412	11
16	415	10
17	418	10
18	421	9
19	424	9
20	427	9
21	429	8
22	431	8
23	434	8
24	436	8
25	438	8
26	440	7
27	442	7
28	444	7
29	446	7
30	448	7
31	450	7
32	452	7
33	454	7
34	455	7
35	457	7
36	459	7
37	461	7
38	463	7
39	465	7
40	467	7
41	469	7
42	470	7
43	472	7
44	475	7
45	477	7

Table 18E. Raw Score to Scale Score Table Grade Span 1: 2, Oral (continued)

Raw Score	Scale Score	Standard Error
46	479	7
47	481	7
48	484	7
49	486	8
50	489	8
51	492	8
52	495	9
53	498	9
54	501	9
55	505	10
56	510	11
57	515	12
58	522	13
59	529	15
60	540	18
61	558	25
62	620	87

Table 19E. Raw Score to Scale Score Table Grade Span 1: K–2, Speaking

Raw Score	Scale Score	Standard Error
0	300	76
1	359	27
2	380	18
3	391	14
4	400	12
5	406	11
6	411	10
7	416	9
8	420	9
9	424	9
10	427	8
11	430	8
12	433	8
13	436	8
14	439	8
15	442	8
16	445	8
17	447	8
18	450	8
19	453	8
20	455	8
21	458	8
22	461	8
23	463	8
24	466	8
25	469	8
26	471	8
27	474	8
28	477	8
29	480	8
30	483	9
31	487	9
32	490	9
33	494	10
34	499	10
35	504	11
36	509	12
37	516	13
38	524	15
39	535	18
40	553	25
41	580	43

Table 20E. Raw Score to Scale Score Table Grade Span 2: 3–5, Speaking

Raw Score	Scale Score	Standard Error
0	310	58
1	349	32
2	373	23
3	387	18
4	397	16
5	405	15
6	411	13
7	417	13
8	422	12
9	427	12
10	432	11
11	436	11
12	440	11
13	444	10
14	447	10
15	451	10
16	455	10
17	458	10
18	461	10
19	465	10
20	468	10
21	472	10
22	475	10
23	479	10
24	483	11
25	486	11
26	490	11
27	495	11
28	499	12
29	504	12
30	508	12
31	514	13
32	519	13
33	525	14
34	532	15
35	539	16
36	548	17
37	557	19
38	569	21
39	585	25
40	611	35
41	635	49

Table 21E. Raw Score to Scale Score Table Grade Span 2: 3–5, Listening

Raw Score	Scale Score	Standard Error
0	310	135
1	310	135
2	310	135
3	310	135
4	310	135
5	310	135
6	377	68
7	405	40
8	420	28
9	431	24
10	442	23
11	452	23
12	462	23
13	473	23
14	484	24
15	495	24
16	508	24
17	520	24
18	533	24
19	548	25
20	565	28
21	587	32
22	620	43
23	630	48

Table 22E. Raw Score to Scale Score Table Grade Span 2: 3–5, Reading

Raw Score	Scale Score	Standard Error
0	300	154
1	300	154
2	300	154
3	300	154
4	300	154
5	300	154
6	300	154
7	300	154
8	383	72
9	408	47
10	423	33
11	434	26
12	443	22
13	450	20
14	457	18
15	463	17
16	469	16
17	475	16
18	480	15
19	486	15
20	491	15
21	497	15
22	502	15
23	508	16
24	514	16
25	521	16
26	528	16
27	535	17
28	542	17
29	551	18
30	560	19
31	571	21
32	584	23
33	601	28
34	630	40
35	660	60

Table 23E. Raw Score to Scale Score Table Grade Span 2: 3–5, Writing

Raw Score	Scale Score	Standard Error
0	270	144
1	270	144
2	270	144
3	270	144
4	270	144
5	270	144
6	282	132
7	342	72
8	366	48
9	383	37
10	395	32
11	405	28
12	415	26
13	423	24
14	431	23
15	438	21
16	445	20
17	451	19
18	458	18
19	464	18
20	470	17
21	475	17
22	481	16
23	486	16
24	492	16
25	498	16
26	504	16
27	510	17
28	517	18
29	525	19
30	533	20
31	543	22
32	554	24
33	569	28
34	589	35
35	621	49
36	680	89
37	680	89

Table 24E. Raw Score to Scale Score Table Grade Span 2: 3–5, Comprehension

Raw Score	Scale Score	Standard Error
0	320	128
1	320	128
2	320	128
3	320	128
4	320	128
5	320	128
6	320	128
7	320	128
8	320	128
9	320	128
10	320	128
11	320	128
12	320	128
13	370	78
14	394	54
15	409	39
16	420	32
17	429	27
18	437	24
19	444	22
20	451	20
21	457	18
22	462	17
23	467	16
24	472	15
25	477	14
26	481	14
27	485	14
28	490	13
29	494	13
30	498	13
31	503	13
32	507	13
33	511	13
34	516	14
35	521	14
36	526	14
37	531	14
38	537	14
39	543	15
40	549	15
41	556	16
42	563	17
43	572	18
44	582	20
45	595	23

Table 24E. Raw Score to Scale Score Table Grade Span 2: 3–5, Comprehension (continued)

Raw Score	Scale Score	Standard Error
46	613	29
47	643	41
48	660	51

Table 25E. Raw Score to Scale Score Table Grade Span 2: 3–5, Oral

Raw Score	Scale Score	Standard Error
0	290	116
1	290	116
2	290	116
3	290	116
4	290	116
5	290	116
6	290	116
7	348	58
8	369	37
9	381	27
10	391	22
11	398	19
12	404	17
13	409	15
14	414	14
15	418	13
16	422	12
17	426	12
18	429	11
19	432	11
20	435	11
21	438	10
22	441	10
23	444	10
24	447	10
25	449	10
26	452	9
27	454	9
28	457	9
29	459	9
30	462	9
31	464	9
32	467	9
33	469	9
34	472	9
35	475	9
36	477	10
37	480	10
38	483	10
39	485	10
40	488	10
41	491	10
42	494	11
43	498	11
44	501	11
45	504	11

Table 25E. Raw Score to Scale Score Table Grade Span 2: 3–5, Oral (continued)

Raw Score	Scale Score	Standard Error
46	508	11
47	512	12
48	516	12
49	520	12
50	524	13
51	529	13
52	533	14
53	539	14
54	544	15
55	550	16
56	557	17
57	565	18
58	574	20
59	584	22
60	597	25
61	614	30
62	638	39
63	680	59
64	680	59

Table 26E. Raw Score to Scale Score Table Grade Span 3: 6–8, Speaking

Raw Score	Scale Score	Standard Error
0	325	59
1	362	33
2	385	23
3	399	19
4	409	17
5	417	15
6	423	14
7	429	13
8	434	12
9	439	12
10	443	11
11	447	11
12	451	11
13	455	11
14	458	11
15	462	10
16	465	10
17	469	10
18	472	10
19	475	10
20	479	10
21	482	10
22	485	10
23	489	11
24	492	11
25	496	11
26	500	11
27	504	11
28	508	12
29	512	12
30	517	12
31	522	13
32	527	13
33	533	14
34	540	15
35	547	16
36	556	18
37	566	20
38	580	23
39	598	28
40	629	40
41	645	48

Table 27E. Raw Score to Scale Score Table Grade Span 3: 6–8, Listening

Raw Score	Scale Score	Standard Error
0	360	64
1	360	64
2	360	64
3	360	64
4	360	64
5	374	51
6	395	38
7	411	31
8	424	28
9	436	25
10	447	24
11	457	24
12	467	23
13	477	23
14	488	24
15	499	25
16	511	26
17	525	28
18	541	32
19	561	37
20	587	41
21	619	43
22	640	46
23	640	46

Table 28E. Raw Score to Scale Score Table Grade Span 3: 6–8, Reading

Raw Score	Scale Score	Standard Error
0	380	101
1	380	101
2	380	101
3	380	101
4	380	101
5	380	101
6	380	101
7	380	101
8	410	70
9	434	46
10	449	33
11	460	26
12	469	22
13	477	20
14	483	18
15	490	17
16	496	16
17	501	16
18	507	16
19	513	16
20	518	16
21	524	16
22	530	16
23	536	16
24	542	16
25	549	17
26	556	17
27	563	18
28	571	18
29	580	19
30	590	20
31	601	22
32	615	26
33	636	34
34	676	58
35	690	68

Table 29E. Raw Score to Scale Score Table Grade Span 3: 6–8, Writing

Raw Score	Scale Score	Standard Error
0	300	128
1	300	128
2	300	128
3	300	128
4	300	128
5	300	128
6	315	112
7	361	67
8	383	45
9	398	35
10	410	29
11	420	26
12	429	24
13	437	22
14	444	21
15	451	20
16	457	19
17	464	19
18	470	18
19	477	18
20	483	18
21	490	18
22	497	18
23	503	18
24	511	18
25	519	19
26	527	20
27	537	22
28	547	23
29	559	24
30	572	26
31	587	28
32	604	31
33	625	34
34	651	34
35	681	38
36	690	44

Table 30E. Raw Score to Scale Score Table Grade Span 3: 6–8, Comprehension

Raw Score	Scale Score	Standard Error
0	360	103
1	360	103
2	360	103
3	360	103
4	360	103
5	360	103
6	360	103
7	360	103
8	360	103
9	360	103
10	360	103
11	360	103
12	360	103
13	375	88
14	403	60
15	420	44
16	432	34
17	441	28
18	449	24
19	456	22
20	462	20
21	468	18
22	473	17
23	478	16
24	483	15
25	487	15
26	492	14
27	496	14
28	500	14
29	504	13
30	508	13
31	513	13
32	517	13
33	521	13
34	525	13
35	530	13
36	535	14
37	539	14
38	544	14
39	550	14
40	555	15
41	561	15
42	567	15
43	574	16
44	581	16
45	590	17

Table 30E. Raw Score to Scale Score Table Grade Span 3: 6–8, Comprehension (continued)

Raw Score	Scale Score	Standard Error
46	599	19
47	611	22
48	627	26
49	655	39
50	680	56

Table 31E. Raw Score to Scale Score Table Grade Span 3: 6–8, Oral

Raw Score	Scale Score	Standard Error
0	310	98
1	310	98
2	310	98
3	310	98
4	310	98
5	310	98
6	310	98
7	349	59
8	370	39
9	383	30
10	394	24
11	402	21
12	409	18
13	415	17
14	420	15
15	425	14
16	429	13
17	433	13
18	436	12
19	440	12
20	443	11
21	446	11
22	449	11
23	452	10
24	455	10
25	458	10
26	460	10
27	463	10
28	466	10
29	468	10
30	471	10
31	473	10
32	476	10
33	478	10
34	481	10
35	484	10
36	486	10
37	489	10
38	491	10
39	494	10
40	497	10
41	499	10
42	502	10
43	505	10
44	508	11
45	511	11

Table 31E. Raw Score to Scale Score Table Grade Span 3: 6–8, Oral (continued)

Raw Score	Scale Score	Standard Error
46	514	11
47	518	11
48	521	12
49	525	12
50	529	13
51	533	13
52	538	14
53	543	14
54	549	15
55	555	16
56	562	17
57	569	19
58	579	20
59	589	23
60	602	25
61	619	29
62	642	35
63	678	50
64	700	63

Table 32E. Raw Score to Scale Score Table Grade Span 4: 9–12, Speaking

Raw Score	Scale Score	Standard Error
0	330	61
1	370	29
2	390	20
3	402	16
4	410	14
5	417	13
6	422	12
7	427	11
8	431	10
9	435	10
10	439	10
11	442	9
12	445	9
13	448	9
14	451	9
15	454	9
16	457	9
17	459	8
18	462	8
19	465	8
20	468	9
21	470	9
22	473	9
23	476	9
24	479	9
25	482	9
26	485	9
27	488	10
28	492	10
29	496	10
30	500	11
31	505	11
32	510	12
33	515	13
34	521	14
35	528	15
36	536	16
37	546	18
38	558	21
39	574	26
40	602	38
41	650	72

Table 33E. Raw Score to Scale Score Table Grade Span 4: 9–12, Listening

Raw Score	Scale Score	Standard Error
0	370	85
1	370	85
2	370	85
3	370	85
4	370	85
5	370	85
6	373	82
7	404	51
8	424	38
9	439	31
10	451	28
11	463	26
12	473	24
13	484	24
14	494	23
15	505	23
16	516	24
17	528	25
18	541	26
19	556	29
20	575	33
21	601	40
22	644	59
23	650	62

Table 34E. Raw Score to Scale Score Table Grade Span 4: 9–12, Reading

Raw Score	Scale Score	Standard Error
0	390	119
1	390	119
2	390	119
3	390	119
4	390	119
5	390	119
6	390	119
7	390	119
8	432	77
9	463	46
10	478	31
11	488	24
12	496	20
13	503	18
14	509	16
15	514	15
16	520	15
17	525	14
18	530	14
19	535	14
20	540	14
21	545	14
22	550	14
23	556	14
24	561	14
25	567	15
26	574	15
27	580	16
28	587	16
29	595	17
30	604	18
31	614	19
32	626	22
33	644	28
34	675	46
35	700	65

Table 35E. Raw Score to Scale Score Table Grade Span 4: 9–12, Writing

Raw Score	Scale Score	Standard Error
0	310	126
1	310	126
2	310	126
3	310	126
4	310	126
5	310	126
6	310	126
7	336	100
8	381	55
9	399	37
10	411	29
11	421	25
12	429	22
13	437	21
14	444	20
15	450	19
16	457	19
17	463	18
18	470	18
19	477	17
20	483	16
21	489	16
22	495	15
23	502	15
24	508	16
25	515	16
26	522	17
27	530	17
28	539	18
29	548	19
30	559	21
31	571	22
32	584	24
33	601	28
34	623	34
35	663	56
36	700	93

Table 36E. Raw Score to Scale Score Table Grade Span 4: 9–12, Comprehension

Raw Score	Scale Score	Standard Error
0	380	97
1	380	97
2	380	97
3	380	97
4	380	97
5	380	97
6	380	97
7	380	97
8	380	97
9	380	97
10	380	97
11	380	97
12	380	97
13	395	83
14	420	57
15	436	42
16	447	33
17	456	28
18	464	24
19	471	22
20	477	20
21	483	18
22	488	17
23	493	16
24	498	16
25	503	15
26	507	15
27	512	15
28	516	14
29	521	14
30	525	14
31	529	14
32	534	14
33	538	14
34	543	14
35	548	14
36	552	14
37	558	14
38	563	15
39	568	15
40	574	15
41	580	16
42	587	16
43	594	17
44	602	18
45	611	19

Table 36E. Raw Score to Scale Score Table Grade Span 4: 9–12, Comprehension (continued)

Raw Score	Scale Score	Standard Error
46	621	20
47	634	23
48	652	29
49	683	45
50	700	58

Table 37E. Raw Score to Scale Score Table Grade Span 4: 9–12, Oral

Raw Score	Scale Score	Standard Error
0	320	97
1	320	97
2	320	97
3	320	97
4	320	97
5	320	97
6	320	97
7	320	97
8	360	58
9	379	38
10	392	28
11	401	22
12	409	19
13	415	17
14	420	15
15	424	14
16	428	13
17	432	12
18	435	11
19	438	11
20	441	10
21	444	10
22	447	10
23	449	9
24	452	9
25	454	9
26	456	9
27	458	9
28	461	9
29	463	8
30	465	8
31	467	8
32	469	8
33	471	8
34	473	8
35	476	9
36	478	9
37	480	9
38	482	9
39	485	9
40	487	9
41	490	9
42	492	10
43	495	10
44	498	10
45	501	10

Table 37E. Raw Score to Scale Score Table Grade Span 4: 9–12, Oral (continued)

Raw Score	Scale Score	Standard Error
46	504	11
47	508	11
48	511	11
49	515	12
50	519	12
51	523	13
52	528	13
53	533	14
54	538	15
55	544	16
56	551	17
57	558	18
58	567	20
59	577	23
60	590	26
61	606	31
62	630	40
63	674	64
64	710	93

Colorado English Language Acquisition Assessment Program

2010 Technical Report

Appendix F: List of Colorado Standards Directly Assessed
by the CELApro

The following tables show the standards in each domain that are directly assessed in the CELApro.

Grades K–2 Listening	
a. Beginning Level	
1. Follow clear one-step directions in 1:1 situations.	
2. Respond nonverbally or in one- or two-word phrases to greetings and requests in 1:1 situations.	
4. Demonstrate comprehension of stories and information by responding nonverbally, e.g., listening, pointing, moving, matching, drawing, gesturing, or with simple word responses.	
5. Understand key words and familiar phrases.	
b. Intermediate Level	
1. Follow clear two- or three-step directions in 1:1 situations.	
2. Respond to familiar social language.	
3. Respond to simple and some complex questions with words or phrases, e.g., either/or, who/how.	
4. Comprehend stories, key concepts of content area information, and oral presentations with contextual support, e.g., graphic organizers, posters, diagrams, charts.	
5. Identify characters and sequence in a story.	
c. Advanced Level	
1. Follow multiple-step directions.	
3. Demonstrate comprehension of read-aloud stories, content area information, and oral presentations with contextual support by paraphrasing, explaining, expressing opinions, etc.	
4. Respond to complex and content-related questions about newly learned information, e.g., compare and contrast.	
5. Identify main idea and details in a story.	
6. Demonstrate understanding of some idiomatic expressions by responding appropriately to such expressions as "hit the road."	

Grades K–2 Speaking	
a. Beginning Level	
1. Use gestures, single words, and simple phrases in greetings, routine conversations, and to communicate needs.	
2. Use both social and academic learned vocabulary in context.	
5. Retell simple stories and personal experiences using key words and phrases.	
6. Respond to simple questions related to immediate context with single words, e.g., yes/no, either/or, basic personal information.	
b. Intermediate Level	
1. Imitate appropriate language in formal and informal settings.	
2. Use vocabulary learned in content area lessons.	
3. Use simple sentences to express needs and ideas about familiar topics in social and academic contexts.	
4. Retell familiar stories and experiences using simple sentences.	
6. Contribute to classroom and small group discussions by responding to and asking simple questions.	

c. Advanced Level
1. Communicate information and feelings clearly in conversations.
2. Contribute to classroom discussions by asking/answering questions, giving opinions, disagreeing, and sharing experiences.
3. Retell, narrate, and paraphrase stories with descriptive detail including characters and setting.
4. Use both formal and informal language, e.g., interviewing, persuasive speech, with attention to grammar, vocabulary, intonation, and pronunciation.
5. Use expanded and descriptive vocabulary related to content areas.

Grades K–2 Reading	
a. Beginning Level	
3. Demonstrate understanding of sound-symbol relationships appropriate to grade level.	
4. Use illustrations and background knowledge to create meaning from read-aloud stories.	
8. Recognize English phonemes that students are already able to hear and produce.	
9. Recognize patterns in writing, e.g., rhyming, word families.	
10. Recognize the difference between letters and numbers.	
b. Intermediate Level	
1. Read patterned language and familiar stories at appropriate grade level.	
2. Have command of sound-symbol relationship and basic rules of word formation.	
5. Use illustrations and background knowledge in combination with more complex vocabulary to create meaning and predict text.	
6. Read and understand the meaning of words, phrases, and short sentences with familiar vocabulary and text independently.	
7. Respond to simple stories read aloud using simple sentences.	
9. Demonstrate an understanding of prediction, main idea, and sequence using key words and simple sentences.	
11. Comprehend key concepts of grade-appropriate content area text supported within an instructional unit.	
c. Advanced Level	
2. Have command of sound-symbol relationship and basic rules of word formation in phrases, simple sentences, and simple texts.	
4. Demonstrate ability to comprehend text and use new information to respond appropriately.	
5. Respond appropriately to questions about main idea and story elements.	
6. Answer questions about inferences, cause and effect, problem solving.	

Grades K–2 Writing	
a. Beginning Level	
7. Label familiar objects and pictures, including key parts of content area objects.	
9. Write pattern sentences with assistance.	
b. Intermediate Level	
1. Demonstrate knowledge of directionality, space between words, and upper-/lower-case letters.	
2. Supply missing words in familiar or framed patterns.	
6. Write simple sentences with spelling approximations and accurate end punctuation.	
9. Write a simple narrative with beginning, middle, and end including sequencing words, e.g., next, then with guidance.	
10. Write a simple descriptive sentence based on observation and learned vocabulary related to content areas.	
c. Advanced Level	
1. Write stories about a series of pictures sequentially and with details.	
3. Write a variety of sentences, e.g., interrogative, declarative, and exclamatory.	
4. Write simple sentences and short paragraphs to demonstrate content area learning.	
5. Write short narrative stories that include the elements of setting, characters, and events.	
8. Communicate in writing using a variety of genres, e.g., expository, descriptive, letters, poetry, autobiography.	
9. Use conventional spelling, capitalization, punctuation, grammar, and sentence structure.	

Grades 3–5 Listening	
a. Beginning level	
1. Follow clear one-step directions in 1:1 situations.	
2. Respond non-verbally or in one- or two-word phrases to greetings and requests in 1:1 situations.	
4. Demonstrate comprehension of stories, information, and academic content by responding non-verbally, e.g., listening, pointing, moving, matching, drawing, and gesturing.	
5. Understand key words and familiar phrases.	
7. Listen for specific purposes and main ideas.	
b. Intermediate Level	
1. Follow clear multiple-step directions in 1:1 situations.	
2. Respond to familiar social language.	
3. Respond to simple and some complex questions with words and phrases, e.g., open-ended, either/or, who/how.	
4. Comprehend stories, key concepts of content area information, and oral presentations with contextual support, e.g., graphic organizers, posters, diagrams, charts.	
5. Identify characters, setting, and sequence in a story.	
c. Advanced Level	
1. Follow multiple-step directions.	
3. Demonstrate comprehension of read-aloud stories, content area information, and oral presentations with contextual support by paraphrasing, explaining, expressing opinions, etc.	
4. Respond to complex and content-related questions about newly learned information, e.g., compare and contrast, describe.	
5. Identify main idea and details in a story.	
6. Demonstrate understanding of some idiomatic expressions by responding appropriately to such expressions as "hit the road."	

Grades 3–5 Speaking	
a. Beginning Level	
1. Use gestures, single words, and simple phrases in greetings, routine conversations, and to communicate needs.	
2. Use both social and academic learned vocabulary in context.	
5. Retell predictable and familiar stories using single words and phrases.	
6. Respond to simple questions related to immediate context with single words, e.g., yes/no, open-ended, either/or, basic personal information.	
b. Intermediate Level	
2. Use key content and descriptive vocabulary.	
3. Communicate ideas using simple sentences in social and academic settings.	
4. Retell familiar stories and identify main idea and some details using phrases and sentences.	
6. Contribute to classroom and small group discussions by responding to and asking simple questions.	
c. Advanced Level	
1. Communicate information, observations, and ideas, and express feelings clearly in conversations.	
3. Contribute to classroom discussions by asking/answering questions, giving opinions, disagreeing.	
4. Retell, narrate, and paraphrase stories with descriptive detail including characters, setting, plot, summary, and analysis.	
5. Use both formal and informal language, e.g., interviewing, persuasive speech, with attention to grammar, vocabulary, intonation, and pronunciation.	
7. Use technical, expanded, and descriptive vocabulary related to content areas.	

Grades 3–5 Reading	
a. Beginning Level	
1. Demonstrate understanding of sound-symbol relationship in context and in own reading.	
3. Recognize simple words, familiar phrases, and simple sentences.	
5. Recognize sight words and content area vocabulary that have been taught in context.	
8. Answer in one- or two-word phrases literal questions about text.	
b. Intermediate Level	
1. Have command of sound-symbol relationship and basic rules of word formation in own reading.	
2. Understand the meaning of words, phrases, and short sentences with familiar vocabulary in text.	
4. Recognize some common root words, prefixes, and suffixes when they are attached to known vocabulary.	
5. Identify basic story elements and make predictions in stories and information using simple sentences.	
6. Use a variety of reading strategies to interpret the meaning of unfamiliar words and gain information from text, print, and media resources, e.g., illustrations, decoding, word attack skills, background knowledge.	
9. Comprehend key concepts of grade-appropriate content area text supported within an instructional unit.	

c. Advanced Level
1. Apply reading strategies to understand more complex text and unfamiliar words, e.g., decoding, previewing, skimming/scanning, problem solving, word attack skills, prior knowledge.
2. Comprehend grade-appropriate content area text with support.
3. Make inferences about information and ideas with reference to features in text.
4. Respond to stories/text by analyzing, summarizing, giving opinions, evaluation, justifying, examining, and explaining.
5. Recognize common root words, prefixes, and suffixes.

Grades 3–5 Writing	
a. Beginning Level	
3. Label familiar objects and pictures including key parts of content area objects.	
6. Write pattern sentences with assistance.	
7. Write about personal experiences with support and using illustrations and simple phrases.	
8. Use familiar words presented and emphasized in content, e.g., literature, math, and science area lessons.	
b. Intermediate Level	
1. Respond to visual or written prompts using phrases and simple sentences.	
3. Use standard word order and conventional spelling, punctuation, and capitalization in basic sentence patterns with support.	
6. Write short descriptive and expository paragraphs for academic communication.	
8. Use transition words to organize writing, e.g., first, after that, in the beginning, later.	
10. Write to express preferences and opinions.	
c. Advanced Level	
1. Use print for social communication, e.g., notes, invitations, e-mail, diary entries.	
2. Respond appropriately to a prompt using narrative, expository, or persuasive writing.	
3. Use a variety of sentence patterns with appropriate tenses and conventions.	
7. Write descriptions and comparisons, including the use of figurative language.	
8. Use conventional spelling, capitalization, punctuation, grammar, and sentence structure.	

Grades 6–8 Listening	
a. Beginning Level	
1. Follow clear one-step directions in 1:1 and group situations.	
2. Respond non-verbally or in one- or two-word phrases to greetings and requests in 1:1 and group situations.	
4. Demonstrate comprehension of stories, information, and academic content by responding non-verbally, e.g., listening, pointing, moving, matching, drawing, and gesturing.	
5. Understand key words, phrases, and simple sentences.	
7. Listen for specific purposes, main ideas, and details.	

b. Intermediate Level
1. Follow clear multiple-step directions in group situations.
2. Respond to social and familiar academic language.
3. Respond to simple and some complex questions with words and phrases, e.g., open-ended, either/or, who/how.
4. Comprehend stories, key concepts of content area information, and oral presentations with contextual support, e.g., graphic organizers, posters, diagrams, charts.
5. Identify story elements, e.g., characters, setting, plot.
c. Advanced Level
1. Follow complex directions involving multiple options and choices.
3. Analyze and evaluate conversations and orally presented stories and content.
4. Comprehend stories and content area concepts at or near grade level with contextual support, e.g., graphic organizers, posters, diagrams, and charts.
5. Respond to complex and content-related questions about newly learned information, e.g., inference, comparison, summarization, point of view, disagreeing.

Grades 6–8 Speaking	
a. Beginning Level	
1. Use gestures, single words, and simple phrases during basic conversations and to communicate needs in social and academic settings.	
2. Use both social and academic learned vocabulary in context.	
5. Retell predictable and familiar stories using single words and phrases.	
6. Respond to simple questions related to immediate context with single words, phrases, e.g., yes/no, either/or, basic personal information.	
7. Ask questions to obtain and clarify information using single words and phrases.	
b. Intermediate Level	
2. Use key content and descriptive vocabulary.	
3. Communicate ideas about a wide range of topics, both social and academic, using simple sentences.	
4. Restate information and identify main idea and some details using sentences, e.g., oral presentations, texts, media, etc.	
6. Contribute to classroom and small group discussions by asking and responding to questions to obtain, clarify, and extend information.	
c. Advanced Level	
1. Communicate information, observations, and ideas, and express feelings clearly in conversations.	
2. Engage in collaborative activities through a variety of student groupings to gather, share, express, and interpret opinions, organize and present information.	
7. Use technical, expanded, and descriptive vocabulary related to content areas.	

Grades 6–8 Reading	
a. Beginning Level	
1. Demonstrate knowledge of sound-symbol relationship in context and in own reading and writing.	
4. Identify and use beginning reading strategies to make text comprehensible and meaningful, e.g., illustrations, graphic organizers, text features, pacing, word attack skills, picture/bilingual dictionaries.	
5. Identify main ideas, details, and sequence of events from modified/visually-supported texts based upon purpose for reading, using non-verbal or one- or two-word phrases.	
b. Intermediate Level	
1. Recognize, decode, and pronounce new vocabulary in context.	
2. Comprehend key concepts of grade-appropriate content area text with support.	
4. Identify and use reading strategies to gain information and make text comprehensible and meaningful, e.g., predicting, questioning, summarizing, self-correcting, evaluating, compare/contrast, print, and media resources.	
5. Identify elements of a story, main ideas, details, and sequence of events from modified/visually-supported texts based upon purpose for reading using simple sentences and paragraphs.	
6. Use grade-appropriate syntax/contextual clues to gain meaning from new vocabulary.	
c. Advanced Level	
1. Use a variety of reading strategies to understand more complex text and unfamiliar words from print and media resources, e.g., paraphrasing, skimming/scanning, problem solving, syntax, summarizing, compare/contrast, and cause/effect.	
2. Comprehend grade-appropriate content area text with support.	
3. Identify elements of a story, main ideas, details, and sequence of events of near or at grade-level text.	
5. Use multiple resources at grade level to draw inferences, conclusions, and generalizations, e.g., glossaries, indexes, graphs, illustrations, headings, subheadings, key vocabulary.	
6. Respond to near or at grade-level text by defending, justifying, supporting inferences, and evaluating.	
9. Demonstrate understanding of some language subtleties, e.g., common idioms, dialect, humor, figurative language.	

Grades 6–8 Writing	
a. Beginning Level	
2. Label objects and illustrations presented in content area lessons.	
3. Write simple sentences and phrases using a model.	
5. Write about personal and group experiences using illustrations, words, and phrases.	
b. Intermediate Level	
1. Use basic sentence patterns with conventional spelling, capitalization, and punctuation.	
7. Write paragraphs and short expository compositions on content area topics, e.g., compare/contrast, cause/effect, problem/solution that include an introductory statement, supporting details and conclusion.	
9. Communicate in writing using a variety of genres, e.g., narrative, content area reports, letter writing, poetry, autobiography.	

c. Advanced Level
1. Respond appropriately to a prompt using narrative, expository, or persuasive writing.
5. Use complex sentence structures with grade-appropriate vocabulary, appropriate syntax, and conventions, e.g., spelling, capitalization, punctuation.
10. Write clear and accurate descriptions and comparisons, including the use of figurative language.

Grades 9–12 Listening	
a. Beginning Level	
1. Follow clear multiple-step directions in 1:1 and group situations.	
2. Respond non-verbally or in one- or two-word phrases to greetings and requests in 1:1 and group situations.	
4. Demonstrate comprehension of stories, information, and academic content by responding non-verbally, e.g., listening, pointing, moving, matching, drawing, and gesturing.	
5. Understand key words, phrases, and simple sentences.	
7. Listen for specific purposes, main ideas, and details.	
b. Intermediate Level	
1. Follow specific multiple-step directions in group situations.	
2. Respond to social and familiar academic language.	
3. Respond to simple and some complex questions with words and phrases, e.g., open-ended, either/or, who/how.	
4. Comprehend stories, key concepts of content area information, and oral presentations with contextual support, e.g., graphic organizers, posters, diagrams, charts.	
5. Identify story elements, e.g., characters, setting, plot, and theme.	
c. Advanced Level	
1. Follow complex directions involving multiple options and choices.	
3. Analyze and evaluate conversations and orally presented stories and content.	
4. Comprehend stories and content area concepts at or grade level with contextual support, e.g., graphic organizers, posters, diagrams, and charts, by analyzing, evaluating, examining, etc.	
5. Respond to increasingly complex and content-related questions about newly learned information, e.g., inference, comparison, summarization, point of view, disagreeing.	
6. Demonstrate understanding of most oral language subtleties, e.g., figurative language, humor, idioms, sarcasm, riddles, slang.	

Grades 9–12 Speaking	
a. Beginning Level	
1. Use gestures, single words, and simple phrases during basic conversations and to communicate needs in social and academic settings.	
2. Use both social and academic learned vocabulary in context.	
5. Restate information from social and academic settings using single words and phrases.	
6. Respond to simple questions related to immediate context with single words, phrases, e.g., yes/no, open-ended, either/or, basic personal information.	

b. Intermediate Level
2. Use key content and descriptive vocabulary.
3. Communicate ideas about a wide range of topics, both social and academic, using simple sentences.
4. Restate information and identify main idea and some details using sentences, e.g., oral presentations, texts, media, etc.
6. Contribute to classroom and small group discussions by asking and responding to questions to obtain, clarify, and extend information.
c. Advanced Level
1. Communicate information, observations, and ideas, and express feelings clearly in conversations.
2. Engage in collaborative activities through a variety of student groupings to gather, share, express, and interpret opinions, discuss, reflect on, organize, analyze, synthesize, and present information.
7. Use technical, expanded, and descriptive vocabulary related to content areas.

Grades 9–12 Reading	
a. Beginning Level	
1. Recognize, decode, and pronounce high frequency words.	
3. Identify and use reading strategies to make text comprehensible and meaningful, e.g., illustrations, graphic organizers, text features, pacing, word attack skills, resource materials.	
5. Identify elements of story, main ideas, details, and sequence of events from modified/visually-supported texts using non-verbal or one- or two- word phrases.	
b. Intermediate Level	
1. Recognize, decode, and pronounce high frequency vocabulary with ease and comprehension.	
2. Comprehend key concepts of grade-appropriate content area text with support.	
4. Identify and use reading strategies to gain information and make text comprehensible and meaningful, e.g., previewing/reviewing, skimming, and identifying related topics and main ideas, print and media resources.	
6. Identify elements of story, main ideas, details, and sequence of events by using context, based upon purpose for reading.	
7. Understand and analyze text for literal and implied meaning.	
c. Advanced Level	
1. Use a variety of reading strategies to comprehend at or near grade-level text and extend personal knowledge from print and media resources, e.g., paraphrasing, previewing/reviewing, skimming/scanning, summarizing, evaluating, compare/contrast, cause/effect, examining.	
2. Comprehend grade-appropriate content area text with support.	
4. Support interpretation of text for literal and implied meaning with reference to features in written text, e.g., vocabulary, facts, sequence, relevance of details, and bias of author.	
6. Demonstrate understanding of most language subtleties, e.g., common idioms, dialect, humor, figurative language.	

Grades 9–12 Writing	
a. Beginning Level	
1. Write simple sentences and phrases using a model.	
4. Write a short personal narrative using simple sentences.	

b. Intermediate Level
3. Use varied sentence patterns with conventional spelling, capitalization, and punctuation.
5. Write, with modeling and support, paragraphs and short expository compositions on content area topics, e.g., compare/contrast, cause/effect, problem/solution that include an introductory statement, supporting details, and a conclusion.
9. Communicate in writing using a variety of genres, e.g., research papers, business letters, editorials, poetry, and autobiography.
c. Advanced Level
3. Respond appropriately to open-ended prompt in narrative, expository, or persuasive writing.
6. Use conventional spelling, capitalization, punctuation, grammar, and syntax.